

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2006, 19:14:26 ; Search time 134 Seconds
(without alignments)
2167.383 Million cell updates/sec

Title: US-09-463-733-1

Sequence: 1 MGENA1RA1FIQWYRRHQ.....VHHD1PTDCESKVIDPKKS 661

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq.21.*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|----------|---------------------|
| 1 | 3315 | 96.3 | 637 | 4 | ABB71489 | Abb71489 Drosophila |
| 2 | 1305 | 37.9 | 722 | 8 | ADN22998 | Adn22998 Bacterial |
| 3 | 1271.5 | 36.9 | 653 | 4 | ABA47250 | Ab47250 Human Pp7 |
| 4 | 1271.5 | 36.9 | 653 | 8 | ADT6673 | Adt6673 Human pro |
| 5 | 1164.5 | 33.8 | 613 | 7 | ADJ95096 | Adj95096 Novel NOV |
| 6 | 1164.5 | 33.8 | 613 | 7 | ADJ95098 | Adj95098 Novel NOV |
| 7 | 581 | 16.9 | 488 | 8 | ADY05391 | Ady05391 Plant ful |
| 8 | 577 | 16.8 | 520 | 4 | ABB63840 | Abb63840 Drosophila |
| 9 | 571 | 16.6 | 500 | 4 | ABG09989 | Abg09989 Novel hum |
| 10 | 570 | 16.6 | 494 | 5 | ABG70122 | Abg70122 Human pre |
| 11 | 570 | 16.6 | 499 | 4 | AAE09722 | Aae09722 Novel cel |
| 12 | 570 | 16.6 | 499 | 7 | ADBE4103 | Adbe4103 Human Pro |
| 13 | 570 | 16.6 | 499 | 8 | ABM82127 | Abm82127 Tumour-as |
| 14 | 569 | 16.5 | 499 | 7 | ADBE4101 | Adbe4101 Rat Prote |
| 15 | 562.5 | 16.3 | 482 | 5 | AAE16244 | Aae16244 Tomato in |
| 16 | 562.5 | 16.3 | 485 | 5 | AAE16236 | Aae16236 Tomato in |
| 17 | 562.5 | 16.3 | 499 | 4 | AAE09733 | Aae09733 Protein p |
| 18 | 560 | 16.3 | 499 | 4 | AAE09732 | Aae09732 Protein p |
| 19 | 556 | 16.2 | 499 | 4 | AAE09732 | Aae09732 Protein p |
| 20 | 519 | 15.1 | 473 | 8 | ADN21438 | Adn21438 Bacterial |
| 21 | 502.5 | 14.6 | 598 | 4 | ABG09988 | Abg09988 Novel hum |
| 22 | 496.5 | 14.4 | 513 | 6 | ABR53882 | Ab53882 Protein s |
| 23 | 479.5 | 13.9 | 513 | 7 | ADK64862 | Adk64862 Disease t |
| 24 | 479.5 | 13.9 | 513 | 7 | ADK64862 | Adk64862 Disease t |

| | | | | | | |
|----|-------|------|-----|---|----------|--------------------|
| 25 | 479.5 | 13.9 | 513 | 8 | ADS43550 | Ad43550 Bacterial |
| 26 | 460.5 | 13.4 | 346 | 4 | ABR59732 | Ab59732 Drosophila |
| 27 | 451 | 13.1 | 244 | 4 | ADS23974 | Ad23974 Bacterial |
| 28 | 435.5 | 12.7 | 323 | 6 | ABP98033 | Abp98033 Protein p |
| 29 | 434.5 | 12.6 | 339 | 8 | ADK87721 | Adk87721 Plant ful |
| 30 | 434 | 12.6 | 318 | 3 | AAG30062 | Aag30062 Arabidops |
| 31 | 434 | 12.6 | 318 | 8 | ADN72465 | Adn72465 Thale cre |
| 32 | 431.5 | 12.5 | 380 | 8 | ADY09284 | Ady09284 Plant ful |
| 33 | 430.5 | 12.5 | 325 | 8 | ADY59840 | Ady59840 Plant pol |
| 34 | 430.5 | 12.5 | 334 | 8 | ADY08660 | Ady08660 Plant ful |
| 35 | 430.5 | 12.5 | 334 | 8 | ADY22333 | Ady22333 Plant ful |
| 36 | 430 | 12.5 | 304 | 3 | AAG30063 | Aag30063 Arabidops |
| 37 | 429.5 | 12.5 | 325 | 8 | ADY57338 | Ady57338 Plant pol |
| 38 | 429.5 | 12.5 | 326 | 8 | ADY59253 | Ady59253 Plant pol |
| 39 | 429 | 12.5 | 338 | 8 | ADY10803 | Ady10803 Plant ful |
| 40 | 429 | 12.5 | 338 | 8 | ADY06935 | Ady06935 Plant ful |
| 41 | 427.5 | 12.4 | 316 | 8 | ADN19921 | Adn19921 Bacterial |
| 42 | 425 | 12.3 | 401 | 8 | ADN23944 | Adn23944 Bacterial |
| 43 | 424.5 | 12.3 | 312 | 3 | AAG39549 | Aag39549 Arabidops |
| 44 | 423 | 12.3 | 290 | 8 | ADK67181 | Adk67181 Plant ful |
| 45 | 423 | 12.3 | 313 | 8 | ADK79431 | Adk79431 Plant ful |

ALIGNMENTS

RESULT 1
ID ABB71489 standard; protein; 637 AA.
XX
AC ABB71489;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 41259.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-UTL-2000; 2000US-00614150.
XX
(PEKE) PE CORP NY.
XX
Venter JC, Adams M, Li PMD, Myers EW;
PI WPI, 2001-656860/75.
XX
N-PSDB; ABL15592.
XX
New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.
XX
Disclosure, SEQ ID NO 41259; 21pp + Sequence Listing; English.
XX
The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL101840-ABL16175) and the encoded proteins (ABR57737-ABR72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 637 AA;

Query Match 96.3%; Score 3315; DB 4; Length 637;
Best Local Similarity 100.0%; Pred. No. 2,6e-304;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

25 MORRCMIOFQNLLEVASBODAELEFEPNDLTKMFOAGRKQYQGSARHVSGLDQDL 84
1 MGRRCMIOFQNLLEVASBODAELEFEPNDLTKMFOAGRKQYQGSARHVSGLDQDL 60
QY 85 VSEFGDIYNAKTELPRKXHDILLIVFRKQGNRLHPYVALILREAKSLKQLENIS 144
61 VSEFGDIYNAKTELPRKXHDILLIVFRKQGNRLHPYVALILREAKSLKQLENIS 120
QY 145 VSTAVSQOYVCGDLHGKLDLVLVHKGGLPSSSPYFNGDFVGRGGLVLLILLS 204
121 VSTAVSQOYVCGDLHGKLDLVLVHKGGLPSSSPYFNGDFVGRGGLVLLILLS 180
QY 205 LYLAFNNAVFLNRGNHDSVMNARYGPIREVBKSPRNKRILAFIDEYRWLPLGSV 264
181 LYLAFNNAVFLNRGNHDSVMNARYGPIREVBKSPRNKRILAFIDEYRWLPLGSV 240
QY 265 SRVLIYHGGSFSDSLDIKSIDRGKYVSIILRPPLTDGSPDXTKMQIFDIMSDPQAT 324
241 SRVLIYHGGSFSDSLDIKSIDRGKYVSIILRPPLTDGSPDXTKMQIFDIMSDPQAT 300
QY 325 MGCVPPTLNGAGVWFGPDVTDNPLQRRRLSYVIRSHCKPNGHGFHMDKITTIPASNY 384
301 MGCVPPTLNGAGVWFGPDVTDNPLQRRRLSYVIRSHCKPNGHGFHMDKITTIPASNY 360
QY 385 YVIGSKGAYIRLNQMLHPFVQYISASQTKLSFKQMGIVSSALKEIAVRMDHD 444
361 YVIGSKGAYIRLNQMLHPFVQYISASQTKLSFKQMGIVSSALKEIAVRMDHD 420
QY 445 ELDEFRKTDPKDSGYISISHMCKVENVTKGLPRLIRDLKLAFTDSQKVVNRTLDL 504
421 ELDEFRKTDPKDSGYISISHMCKVENVTKGLPRLIRDLKLAFTDSQKVVNRTLDL 480
QY 505 LDDVTLLEABDQMSWMDALYNKASIVAFNIIIDNSGETLDFEFAIDLVAMHG 564
481 LDDVTLLEABDQMSWMDALYNKASIVAFNIIIDNSGETLDFEFAIDLVAMHG 540
QY 565 AVSKAEMLEKCRMDLNGDKVDLNEFLAFLSLDLHKKQDENIRRSSTGSPVAKTA 624
541 AVSKAEMLEKCRMDLNGDKVDLNEFLAFLSLDLHKKQDENIRRSSTGSPVAKTA 600
QY 625 TDPVTLADKISKNTLVESHIDPDTCESKVIDPKKS 661
601 TDPVTLADKISKNTLVESHIDPDTCESKVIDPKKS 637
DB

RESULT 2
ID ADN22998
ADN22998 standard; protein; 722 AA.

AC ADN22998;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #5651.
XX
KM Recombinant DNA construct; transformed plant; improved plant property;
KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KM pathogen tolerance; pest tolerance; plant disease resistance;
KM cell cycle pathway modification; plant growth regulator;
KM homologous recombination; seed oil yield; protein yield; carbohydrate;
KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KM bacterial polypeptide.

OS Bacteria.
XX
PN US2003233675-A1.

XX 18-DEC-2003.
PD 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
XX (HINK/) HINK G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS,
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX
XX Claim 1, SEQ ID NO 5651, 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX homologous recombination, modified seed oil or protein yield and/or
XX content, improved yield by modification of carbohydrate, nitrogen or
XX phosphorus use and/or uptake, by modification of photosynthesis or by
XX providing improved plant growth and development under at least one stress
XX condition, improved lignin production or improved galactomannan
XX production. This sequence represents a bacterial polypeptide used in the
XX scope of the invention. Note: The sequence data for this patent did not
XX form part of the printed specification but was obtained in electronic
XX format from USPTO at seqdata.uspto.gov/sequence.html.

XX
SQ Sequence 722 AA;

Query Match 37.9%; Score 1305; DB 8; Length 722;
Best Local Similarity 41.6%; Pred. No. 9,6e-114;
Matches 266; Conservative 117; Mismatches 201; Indels 56; Gaps 13;

QY 4 NAIRAIPIQKWRHQAEREMQRCONQIFQNLLEVASBODAE-----LYKFNLDLK 57
DB 90 SITKSAIILQKWRHQAEREMQRCONQIFQNLLEVASBODAE-----LYKFNLDLK 149
QY 58 HMPQAGRKQYQGSARHVSGLDQDLKXRVWTLVDFADVIR 97
DB 150 AMAERNGKQGVENGNHSPMLSHYAKPSLMDSEGETVKMLLEPTSDPTNVIDRNYK 209
QY 98 ---LPIRKHNDILLIVFRKQGNRLHPYVALILREAKSLKQLENISPVSTAVSQOYT 154
DB 210 TSLSPIDPKQVAKMTBAF--KYNKVLHPYVLMILREAKIRKAMPVSRISTISNQYT 267
QY 155 VCGDLHGKLDLVLVHKGGLPSSSPYFNGDFVGRGGLVLLILLSLYLAFNNAV 214
DB 266 ICGDLHGKFDLCITLKNKGYSDNPIYFNGDFVGRGGLVLLILLSLYLAFNNAV 327
QY 215 LNRGNHDSVMNARYGPIREVBKSPRNKRILAFIDEYRWLPLGSVLSRVLIYHGGF 274
DB 328 LNRGNHDSVMNARYGPIREVBKSPRNKRILAFIDEYRWLPLGSVLSRVLIYHGGF 387

| | | | |
|----------|---|--|-----|
| Oy | 275 | NSSTSLDLILKSIDRGKVVYILAPPLDSEPLDK-----TEWQQLFDLIMSDPQATMGC | 327 |
| Db | 388 | SDQTEVSKLIDKLPRIHFGSVLRRPVNKKMESEKENSANVVDENKQMLDIMSDPKKNGC | 447 |
| Oy | 328 | VPNTLRGAGVWMPDPDTPDNLORHRLSYVIRSHCEKPNGHIEFHNDKILITIFASNYAI | 387 |
| Db | 448 | WPNVFRGGSGSYGADITLASFLEKHGRLLVRSHECKFEGEYFHNNTCLTFPSASNYET | 507 |
| Oy | 388 | GSNKAAYLR-LNNQLMPHFVQYIISAASQTRLS-FQRMGIYESSALKEILAVMRDHRDE | 445 |
| Db | 508 | GSNRGAYVFIKSKQPHFVQYV--ASKTRKSTLLEERLGVVBSAVKEIKELTSLSFHTD | 565 |
| Oy | 446 | LEDERKNDPKDSGYSTISGHCKCMENNYKLGIPMYLLDNDKLAP-GTDSQKVVNNTLDL | 504 |
| Db | 566 | LOKEBIRNDIEKSGKLPILKMSDCVRRITGLNIPWALAPKATLSBDGYVYKEDRRI | 625 |
| Oy | 505 | LDTVDILKEADGMSVMDALVYANKASLVAFINIIDADNGEITLDEFEETAILDLVAHMPG | 564 |
| Db | 626 | AQVGTHNQEKD---IVESLYRHKSTLETLLFRMDQDNGQVSMKEFIDACEVL-----G | 677 |
| Oy | 555 | AYSXKAM-----LEKCRMDLNGDGKVDINEFLAERLSD | 599 |
| Db | 678 | KYTKRPLQTDYISQIAESIDFMNDGFIIDNELLEAFRLVD | 717 |
| RESULT 3 | | | |
| AA | 47250 | | |
| ID | AA847250 | standard; protein; 653 AA. | |
| AC | AA847250; | | |
| XX | 18-JUL-2001 | (first entry) | |
| DE | Human PP7. | | |
| KW | Mitochondrial calcium-binding protein; p95.6/YNS2; calcium; | | |
| XX | drug screening; sequence extraction; PP7; calcium binding motif. | | |
| OS | Homo sapiens. | | |
| PN | WO200130830-A2. | | |
| PD | 03-MAY-2001. | | |
| XX | | | |
| PF | 26-OCT-2000; 2000MO-US029787. | | |
| XX | | | |
| PR | 26-OCT-1999; 99US-00427867. | | |
| XX | 26-OCT-1999; 99US-00427895. | | |
| PR | 07-SEP-2000; 2000US-00657253. | | |
| XX | | | |
| PA | (MITO-) MITOKOR. | | |
| PI | Murphy AN, Fahy ED, Wiley SE, Clevenger W; | | |
| XX | | | |
| DR | WPI: 2001-308616/32. | | |
| XX | N-PSDB; AAC85804. | | |
| PT | New Caenorhabditis elegans mitochondrial calcium-binding protein, | | |
| XX | p95.6/YNS2, identified by protein motif database searching, useful for | | |
| PT | regulating intracellular calcium ions and in drug screening assays. | | |
| XX | | | |
| PS | Example 2; Fig 13; 176pp; English. | | |
| XX | | | |
| CC | This sequence shows a protein containing calmodulin/calcium binding | | |
| CC | motifs which was identified using the method of the invention. The method | | |
| CC | comprises a computer system for extracting information from biological | | |
| CC | sequence databases, which receives a motif that defines a desired | | |
| CC | polypeptide sequence. The computer searches the database based on a query | | |
| CC | formulated from the motif to obtain a set of polypeptide sequences which | | |
| CC | are then filtered and a report of the filtered set of polypeptide | | |
| CC | sequence information is then displayed. The computer system is useful for | | |
| CC | identifying recombinant C.elegans p95.6/YNS2 and other mitochondrial | | |

| CC | protein involved in binding, transport or other regulation of |
|-----------------------|---|
| CC | intracellular calcium which provides targets for screening assays that |
| CC | identify agents capable of altering a mitochondrial activity |
| xx | Sequence 653 AA; |
| Query Match | 36.9%; Score 1271.5; DB 4; Length 653; |
| Best Local Similarity | 40.3%; Pred. No. 1.2e-110; |
| Matches 258; | Conservative 129; Mismatches 204; Indels 49; Gaps 11; |
| QY | 3 ENAIRAIFIOKWRBHOAREMORRCNMOIFONLEVASRDOAELEKFPNDIKKHPQA 62 |
| DB | 14 DTSIAAIIIIINMRYGKARKLAKQHYALITIPQISRYADEQOQULSTFSPMLKNTHTI 73 |
| QY | 63 AGKNOYOGSAHVAVLDDKD--DLVE--BFGDIYNA-KIELEPIKKNHIDLLIDVFRKRG 117 |
| DB | 74 HKSELIELRNGSLSEQOMRDRMDVYDSIDVPDSYNGSRRLQEPPLCTDIDLLLEAFKRGQ- 132 |
| QY | 118 NRIHPKTYALLIREAAKSLKQLEPNISPVSTRASQOQTVCCGDHSGKLDLVLVHKGLPS 177 |
| DB | 133 -ILHAHYLVLELPETKIKVLKQMPNPTIIGTSPSKERTVTCGLHKKLDLPLIPIKNGLPS 191 |
| QY | 178 SSNTYVNGDPVDGKRGLEVLTLLELYLAPAAVAVLNGNHSYMMARVYGTIREVES 237 |
| DB | 192 ERNPVYVNGDVPDVGKNSIBILMLCVFLYVPDNLHANGNHDPMNNLRVYGTKEILH 251 |
| QY | 238 KYPRNHRKRIIAFIDEVYRWMLPLGSLVLSRVLYVIGGFSDSITLDLKSIDRGKVSILRP 297 |
| DB | 252 KYKLGKRIIQLIEEPFAMLPIGITVNDLIVHIGGISFTDMLNLKRVBRNKKSLIIP 311 |
| QY | 298 PL-----TDGEP---LDKTEMOQIFDIMSDDQATYGCV 328 |
| DB | 312 PTEINRBDHDTGSKNKKGVTFMAHGRILKTNGSPTEHLTBHMEWQIDILMSDPRGKNGCF 371 |
| QY | 329 PNTLRGAGWGFDPVDTNFLQRRLSLVYISHSCEKPGNHPMDNKKITTFPSASNTYALG 388 |
| DB | 372 PNTRGGGCYGPGPVTSKILNKYQKMLIRSHCEKPGGYELCHDGKVVITFSAANTYBEG 431 |
| QY | 389 SNKGAVYIRLNQMLPHFVQYISASQTKRSLF---KQRMGVSSALKEILAVRRDRHDE 445 |
| DB | 432 SNKGAVYILGCSGTTTFRFFQY---QVTKATCFQRLRVQDVTMSAKILIRERTSKSD 487 |
| QY | 446 LEDEFKRYKPDQSGYISISHWKVMEVNTYKLGLEPWRLLRDLKALPGTDSOKVNNYRTLDL 505 |
| DB | 488 LTRAFQGLDRHKSGKSLGVSQWAFCEBNILGLNLPMSLSNLVNIIDQNGVEYVSSPONI 547 |
| QY | 506 DTD-VILIEABDMSVMDALYANKASIVAFINIIDADNGSEITLDEBPTAIDLVAHMPG 564 |
| DB | 548 RIEKPVQEAHS---TLVETLYRYRSDLEIIFNAIDTHSGILSVBEPRAWKLFSSHYNV 604 |
| QY | 565 AYSKAEMLKCRMDLNGDKQVLDNEBLEAFRLSDLRKE 604 |
| DB | 605 HIDSQVKNKLINIMDLNKGSDIDNEPLKAFYV--VHRYE 642 |
| RESULT 4 | |
| ADT66673 | |
| ID | ADT66673 standard; protein; 653 AA. |
| XX | ADT66673; |
| AC | |
| DT | 16-DEC-2004 (first entry) |
| XX | |
| DE | Human protein phosphatase, EF hands-1 protein. |
| XX | |
| KM | pain regulation; screening; Lu2P; PBP-19; phosphatidylinositol synthase; |
| KM | valosin; interleukin 6 receptor subunit beta; aspartate aminotransferase; |
| KM | neuronal immediate early gene; heat shock protein 27; HSC70; calmodulin; |
| KM | syntaxin binding protein 1; epliceosoma protein SAP 155; neurodop 1; |
| KM | bamacan; leukotriene A4 hydrolase; chondromodulin 1; |
| KM | Mse4/phosphatidylinositol-4-phosphate 5 kinase; 26S proteasome UB p112; |
| KM | UB Z; UB p44.5; ingensin; 3-hydroxy-3-methylglutaryl-Co A synthase; |
| KM | phosphoglycerate kinase; RNA polymerase II TF SIII p18 UB; |

| | | |
|----|-------------|------------------|
| PR | 05-NOV-2001 | 2001US-0338656P. |
| PR | 06-NOV-2001 | 2001US-0330722P. |
| PR | 09-NOV-2001 | 2001US-0348283P. |
| PR | 15-NOV-2001 | 2001US-0335610P. |
| PR | 16-NOV-2001 | 2001US-0338543P. |
| PR | 20-NOV-2001 | 2001US-0331630P. |
| PR | 20-NOV-2001 | 2001US-0331641P. |
| PR | 21-NOV-2001 | 2001US-0332152P. |
| PR | 27-NOV-2001 | 2001US-0333461P. |
| PR | 28-NOV-2001 | 2001US-0333912P. |
| PR | 28-NOV-2001 | 2001US-0334027P. |
| PR | 29-NOV-2001 | 2001US-0334300P. |
| PR | 30-NOV-2001 | 2001US-0334421P. |
| PR | 30-NOV-2001 | 2001US-0334526P. |
| PR | 04-DEC-2001 | 2001US-0336576P. |
| PR | 04-DEC-2001 | 2001US-0336644P. |
| PR | 07-DEC-2001 | 2001US-0338314P. |
| PR | 07-DEC-2001 | 2001US-0338390P. |
| PR | 10-DEC-2001 | 2001US-0339006P. |
| PR | 10-DEC-2001 | 2001US-0339008P. |
| PR | 11-DEC-2001 | 2001US-0339280P. |
| PR | 01-FEB-2002 | 2002US-0353286P. |
| PR | 01-FEB-2002 | 2002US-0353288P. |
| PR | 04-FEB-2002 | 2002US-0354392P. |
| PR | 04-FEB-2002 | 2002US-0354393P. |
| PR | 04-FEB-2002 | 2002US-0354409P. |
| PR | 27-FEB-2002 | 2002US-0355944P. |
| PR | 27-FEB-2002 | 2002US-0360148P. |
| PR | 05-MAR-2002 | 2002US-0361790P. |
| PR | 05-MAR-2002 | 2002US-0361833P. |
| PR | 05-MAR-2002 | 2002US-0361925P. |
| PR | 05-MAR-2002 | 2002US-0362230P. |
| PR | 05-MAR-2002 | 2002US-0362655P. |
| PR | 13-MAR-2002 | 2002US-0364000P. |
| PR | 13-MAR-2002 | 2002US-0364181P. |
| PR | 13-MAR-2002 | 2002US-0364182P. |
| PR | 13-MAR-2002 | 2002US-0364197P. |
| PR | 13-MAR-2002 | 2002US-0364227P. |
| PR | 17-MAY-2002 | 2002US-0381621P. |
| PR | 17-MAY-2002 | 2002US-0381621P. |
| PR | 28-MAY-2002 | 2002US-0383675P. |
| PR | 17-JUL-2002 | 2002US-0396703P. |
| PR | 06-AUG-2002 | 2002US-0401552P. |
| PR | 07-AUG-2002 | 2002US-0401594P. |
| PR | 07-AUG-2002 | 2002US-0401787P. |
| PR | 15-AUG-2002 | 2002US-0403619P. |
| PR | 20-AUG-2002 | 2002US-0404821P. |
| PR | 23-AUG-2002 | 2002US-0405388P. |
| PR | 23-AUG-2002 | 2002US-0405402P. |
| PR | 23-AUG-2002 | 2002US-0405496P. |
| PR | 23-AUG-2002 | 2002US-0405611P. |
| PR | 26-AUG-2002 | 2002US-0406125P. |
| PR | 04-NOV-2002 | 2002US-00287226. |

| ID | ADJ95098 | standard; protein; 613 AA. |
|----|---|----------------------------|
| XX | ADJ95098; | |
| XX | 06-MAY-2004 | (first entry) |
| XX | Novel NOVX protein sequence #163. | |
| XX | antidiabetic; anorectic; cardiant; hypotensive; antiarteriosclerotic; | |
| KW | anorectic; vitamin; antibiotic; antifungal; procoagulant; noctropic; | |
| KW | antiparkinsonian; anticonvulsant; osteopathic; | |
| KW | antiarthritic; antiinflammatory; dermatological; antisthmatic; | |
| KW | antileptic; gene therapy; metabolic disorder; diabetes; obesity; | |
| KW | infectious disease; anorexia; cancer; cardiovascular disease; | |
| KW | hyperextension; atherosclerosis; neurodegenerative disorder; | |
| KW | Alzheimer's disease; Parkinson's disease; epilepsy; immune disorder; | |
| KW | osteoarthritis; hemopoietic disorder; inflammatory skin disorder; | |
| KW | asthma; dyslipidemia; neurogenesis; cell differentiation; | |
| KW | cell proliferation; hematology; wound healing; angiogenesis; | |
| KW | chromosome mapping; tissue typing; pharmacogenomic. | |
| XX | Homo sapiens. | |
| OS | | |
| XX | MO2003040325-A2. | |
| PN | | |
| XX | 15-MAY-2003. | |
| XX | 05-NOV-2002; 2002WO-US035464. | |
| XX | | |
| PR | 05-NOV-2001; 2001US-0338626P. | |
| PR | 06-NOV-2001; 2001US-0333072P. | |
| PR | 09-NOV-2001; 2001US-0348283P. | |
| PR | 15-NOV-2001; 2001US-0335610P. | |
| PR | 16-NOV-2001; 2001US-0338543P. | |
| PR | 20-NOV-2001; 2001US-0331630P. | |
| PR | 20-NOV-2001; 2001US-0331641P. | |
| PR | 21-NOV-2001; 2001US-0332152P. | |
| PR | 27-NOV-2001; 2001US-0333461P. | |
| PR | 28-NOV-2001; 2001US-0333912P. | |
| PR | 29-NOV-2001; 2001US-0334027P. | |
| PR | 29-NOV-2001; 2001US-0334300P. | |
| PR | 30-NOV-2001; 2001US-0334421P. | |
| PR | 30-NOV-2001; 2001US-0334525P. | |
| PR | 04-DEC-2001; 2001US-0335767P. | |
| PR | 04-DEC-2001; 2001US-0336664P. | |
| PR | 07-DEC-2001; 2001US-0338314P. | |
| PR | 07-DEC-2001; 2001US-0338390P. | |
| PR | 10-DEC-2001; 2001US-0339069P. | |
| PR | 10-DEC-2001; 2001US-0339088P. | |
| PR | 11-DEC-2001; 2001US-0339286P. | |
| PR | 01-FEB-2002; 2002US-0335280P. | |
| PR | 01-FEB-2002; 2002US-0353288P. | |
| PR | 04-FEB-2002; 2002US-0354392P. | |
| PR | 04-FEB-2002; 2002US-0354393P. | |
| PR | 04-FEB-2002; 2002US-0354403P. | |
| PR | 27-FEB-2002; 2002US-0359944P. | |
| PR | 27-FEB-2002; 2002US-0360148P. | |
| PR | 05-MAR-2002; 2002US-0361790P. | |
| PR | 05-MAR-2002; 2002US-0361833P. | |
| PR | 05-MAR-2002; 2002US-0361925P. | |
| PR | 05-MAR-2002; 2002US-0362230P. | |
| PR | 05-MAR-2002; 2002US-0362625P. | |
| PR | 13-MAR-2002; 2002US-0364000P. | |
| PR | 13-MAR-2002; 2002US-0364181P. | |
| PR | 13-MAR-2002; 2002US-0364182P. | |
| PR | 13-MAR-2002; 2002US-0364197P. | |
| PR | 13-MAR-2002; 2002US-0364227P. | |
| PR | 17-MAY-2002; 2002US-0381621P. | |
| PR | 28-MAY-2002; 2002US-0386757P. | |
| PR | 17-JUL-2002; 2002US-0396703P. | |

06-AUG-2002 2002US-0401552P.
07-AUG-2002 2002US-0401584P.
07-AUG-2002 2002US-0401787P.
15-AUG-2002 2002US-0403619P.
20-AUG-2002 2002US-0404821P.
23-AUG-2002 2002US-0405368P.
23-AUG-2002 2002US-0405402P.
23-AUG-2002 2002US-0405496P.
23-AUG-2002 2002US-0405631P.
26-AUG-2002 2002US-0406125P.
04-NOV-2002 2002US-00287226.
(CURA-) CURAGEN CORP.
Agyee ML, Alabrook JP, Beirhs C, Boldog FL, Burgess CE, Chant JS,
Chaudhuri A, Dipippo VA, Eisinger SR, Eisen A, Ellerman K,
Ganggoli EA, Gorman L, Gerlach VL, Ji W, Kekuda R, Khramtsov NV,
Li L, Malyanark UM, Macdonald JR, Mezes PS, Miller CE, Millet I,
Ooi CE, Ott T, Padigara M, Patutarajan M, Raselli L, Rieger DK,
Rottenberg MR, Shenoy SG, Spaderna SK, Spytek KA, Taupier RJ,
Vernet CAM, Zeehunen BD, Zhong M;
MPI; 2003-441551/41.
N-PSDB; ADJ95097.
New isolated NOXY polypeptides and polynucleotides, useful for
preventing, diagnosing or treating NOXY-associated disorders, e.g.
osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
asthma, or infections.
claim 1; SEQ ID NO 326; 800pp; English.

The invention relates to novel isolated polypeptides, mature forms of these, or a sequence that is at least 95 % identical to, or having one or more conservative amino acid substitutions in the polypeptides. The polypeptides, nucleic acid molecules and antibodies are useful in the manufacture of a medicament for treating a syndrome associated with a human disease, preferably a NOXV-associated disorder. The nucleic acid molecules, polypeptides and antibodies are useful for treating, preventing or diagnosing diseases such as metabolic disorders, diabetes, obesity, infectious diseases (viral, bacterial, fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular diseases (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune disorders (osteoarthritis), hematopoietic disorders, inflammatory skin disorders, asthma, and various dyslipidemias. The nucleic acids and polypeptides may also be used as targets for the identification of small molecules that modulate or inhibit e.g. neurogenesis, cell differentiation, cell proliferation, hematopoiesis, wound healing and angiogenesis, in gene therapy, in generation of antibodies that bind immunospecifically to NOXV substances for use in therapeutic or diagnostic methods. The nucleic acids are further used as hybridization probes, in chromosome mapping, tissue typing, preventive medicine, and pharmacogenomics. This sequence corresponds to one of the NOXV polypeptides of the invention.

SQ Sequence 613 AA;

| | | | | |
|-----------------------|-------------------|---------------------|------------|-------------|
| Query Match | 33.8%; | Score 1164.5; | DB 7; | Length 613; |
| Best Local Similarity | 39.3%; | Pred. No. 1.5e-100; | | |
| Matches 247; | Conservative 115; | Mismatches 200; | Indels 67; | Gaps 13 |

```

Oy 3 ENIIRAIIFOKYRHHQARRMORCMNOIROMTEVMSBEOQAEUYEFNDLKHMPQA 62
Oy  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 14 DTSLRPAALIIQNNYRKYARKLRQHVALTIQTSIEVADBOQOMQLSFFSPMLANTHI 73

Oy 63 AGKNOYQOSAHVSUDDKO--DLYE--EFGQIVYA-KLELPRIKNIHDLIDVPRKKRG 11
Db 74 HKESLEIRKNOSESBOEDMRDMDYDSDIVPSTYSGPRLOPFLCTDIDLLBFAKSOQ- 13
Oy 116 NRIHPKYVALIREAAKSLIKOLPNISPVSTAVSQVTVCGDIHGKLDLIDLVLIRKNOJPS 177
Db 133 -IIIAHYVLEVFETKCVLKOMPNTFHIOTSSEKVTIIICGDIHGKLDLIDFLIPKNOJPS 191

```


PN WO2001/1042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX MPI: 2001-656860/75.
DR N-PDSB; ABL07943.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX
PS Disclosure; SEQ ID NO 18312; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB01840-AB146175) and the encoded proteins (AB57737-
XX AB872072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 520 AA;

| | | | | |
|-----------------------|-------|---|-----------------------------------|----------------|
| Query Match | 16.8% | Score 577 | DB 4 | Length 520 |
| Best Local Similarity | 32.1% | Pred. No. 5.5e-45 | | |
| Matches | 136 | Conservative | 83 | Mismatches 153 |
| | | | Indels | 52 |
| | | | Gaps | 15 |
| QY | 4 | NAIRA-AIFIOKMYRHOARREM---- | GRRCMIOIFOMLEVASBODQAEUYKFPNDLIK- | 57 |
| DB | 108 | SAVKADPAVLKGYRRAAAMHSLGKFKQALCDFFVAKCR-PNDQAKLKFTKCKIYVM | | 166 |
| QY | 58 | -----HMQAAGRKQYOGSAHVSYLD-KDVLVEFGDLYNAKTELEPIRKNHIDL | | 107 |
| DB | 167 | RAFERAIAVDKPEKT-LSEMYSDMENTITBEDYKCPQLED-----GKTLKFMKE---- | | 215 |
| QY | 108 | LIDVERKRGKNGELHREYVALILREBAKSLKQLEPNISPVSTAVSOQVTCGDHIGKLDLL | | 167 |
| DB | 216 | LMENHKAOK--RLHKKFAKILICEIDITYRAPSIVDITVPREBKFTICGDHGOFPYDLM | | 273 |
| QY | 168 | VVLHNGELPSSNPVFNQDFVDRGKRGLEVLILLSLVLAEPNAVFLNRGNHDSVMNA | | 227 |
| DB | 274 | NIFEINGLPSEKQPYLFNQDFVDRGSFSVBCIFTLFGFLLLPNHFFLARGNHESINMQ | | 333 |
| QY | 228 | RYGPIREYBSKYPRHMKRIALFIDRYRLPLGSVLNSVVLVHGCF--SDSTSLDIKS | | 285 |
| DB | 334 | MYGFGEVYAKT-----SMAADIFTQVFNALPLCHCNOKILVMHGGLFSTBTVDTHIRK | | 390 |
| QY | 286 | IDRGKYSILRDEPLTDGBPLDKTEWQOILFDIMMSDPQATMGVCVPTLIRGAGVFGPDVTD | | 345 |
| DB | 391 | IERN-----CQPR-----EBSGLMELMSDPQCMMS-LGGSKRGVGIQFGRDYTE | | 434 |
| QY | 346 | NFLQHRRLSYVRSHCEKPNGHFPMHNDKITITFASANTYTAIGSNKAYIRL-NNQLMPH | | 404 |
| DB | 435 | KCKCKNNNDIYIIRSHBVDMGYEVAHNGKCIIVFSAIPNYCDTMGNMGAFITITGNLKEN | | 494 |
| QY | 405 | FVOY 408 | | |
| DB | 495 | YKSF 498 | | |

| | | |
|----------|----|---|
| ABG09989 | ID | ABG09989 standard; protein; 500 AA. |
| XX | AC | ABG09989; |
| XX | DT | 13-FEB-2002 (first entry) |
| XX | DE | Novel human diagnostic protein #9980. |
| XX | KW | Human; chromosome mapping; gene mapping; gene therapy; forensic; |
| XX | KW | food supplement; medical imaging; diagnostic; genetic disorder. |
| XX | OS | Homo sapiens. |
| XX | PM | MO200175067-A2. |
| XX | PD | 11-OCT-2001. |
| XX | PP | 30-MAR-2001; 2001MO-US008631. |
| XX | PR | 31-MAR-2000; 2000US-00540217. |
| XX | PR | 23-AUG-2000; 2000US-00649167. |
| XX | PA | (HYSR-) HYSEQ INC. |
| XX | PI | Drmanc RT, Liu C, Tang YT; |
| XX | DR | WPI, 2001-639362/73. |
| XX | DR | N-PSDB; AAS74176. |
| XX | PT | New isolated polynucleotide and encoded polypeptides, useful in |
| XX | PT | diagnostics, forensics, gene mapping, identification of mutations |
| XX | PT | responsible for genetic disorders or other traits and to assess |
| XX | PT | biodiversity. |
| XX | PS | Claim 20; SEQ ID NO 40348; 103pp; English. |
| XX | CC | The invention relates to isolated polynucleotide (I) and polypeptide (II) |
| XX | CC | sequences. (I) is useful as hybridisation probes, polymerase chain |
| XX | CC | reaction (PCR) primers, oligomers, and for chromosome and gene mapping, |
| XX | CC | and in recombinant production of (II). The polynucleotides are also used |
| XX | CC | in diagnostics as expressed sequence tags for identifying expressed |
| XX | CC | genes. (I) is useful in gene therapy techniques to restore normal |
| XX | CC | activity of (II) or to treat disease states involving (II). (II) is |
| XX | CC | useful for generating antibodies against it, detecting or quantitating a |
| XX | CC | polypeptide in tissue, as molecular weight markers and as a food |
| XX | CC | supplement. (II) and its binding partners are useful in medical imaging |
| XX | CC | of sites expressing (II). (I) and (II) are useful for treating disorders |
| XX | CC | involving aberrant protein expression or biological activity. The |
| XX | CC | polynucleotide and polynucleotide sequences have applications in |
| XX | CC | diagnostics, forensics, gene mapping, identification of mutations |
| XX | CC | responsible for genetic disorders or other traits to assess biodiversity |
| XX | CC | and to produce other types of data and products dependent on DNA and |
| XX | CC | amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic |
| XX | CC | amino acid sequences of the invention. Note: The sequence data for this |
| XX | CC | patent did not appear in the printed specification, but was obtained in |
| XX | CC | electronic format directly from WIPO at |
| XX | CC | ftp.wipo.int/pub/published_pct_sequences |
| XX | CC | Sequence 500 AA; |

| | | | | | |
|-----------------------|---|--------------|--------------------|------------|--------------------------|
| Query Match | | 16.6% | Score 571; | DB 4; | Length 500; |
| Beet Local Similarity | | 34.8% | Pred. No. 1.9e-44; | | |
| Matches | 143; | Conservative | 73; | Mismatches | 153; Indels 42; Gaps 16; |
| QY | 11 FIKKKRRHQAARRRQR--RCNMWIFQNLEYSAS--BDDAELEYKFFNDILKHHPEQAAGRKN | 67 | | | |
| Dd | 99 YIMGYTRRAAANMALGKFPAALRDVEITVVKYKPHDKAMKYQECSNKIVK--PKAFER-- | 154 | | | |
| QY | 68 QYGSAHY-STLDDDD---DLVERFG--DIVNAKIEPIRKNIHTDILLIDVRFKRGGRRL | 120 | | | |
| Dd | 155 AIAAGEHRKSVDSDIADISMTIEBYSXPKLEDGGVTISFMKE---LMQWKOK--KL | 208 | | | |

| | | | |
|-----------|--------------------------------------|---|-----|
| Qy | 121 | HPKVALIILEAAKSLKQLPENISPVSTAAVSQGVGADLHGLDLDLVVLRHNGLPSSN | 180 |
| Dy | 209 | HRKCAVYLIVQVKEVLSKSLSTVETTLTKTEKTIYCGDTHQFPYDLMLFELANGLPSTN | 266 |
| Qy | 181 | PVYVNGDPVNRGRKRGLEVLNLLSLYLAFPNVFLRKGHDSVNNARVGFIREVESKYP | 240 |
| Dy | 269 | PYLINQGFVDRGSPFSVSEVILLTGKFLPLRPHNHLRNGHLEIDNNQIYGFEGVAKATY | 322 |
| Qy | 241 | RNHKRIIAFLDEVYRMLPLGASVNSRVLLVHGQF--SDSTLIDLKSIDRGKYSILRRP | 296 |
| Dy | 329 | ---AQMVELFSEVFEWMLPLAQCCINGKVLIMHGGLFSEDGVTLDDIRIKTERNR-----QRP | 380 |
| Qy | 299 | LTOSERLDKTEWQOQIPIDIMSDPQATMGCVPTLTLGAGVGRPRTDTPFLDRHLSYIR | 350 |
| Dy | 381 | --DSGF-----CDLWMSDPQPNQ--RSISKGGVTCQFPRDVTAKFLERNMDLYIR | 422 |
| Qy | 359 | SHECKRPFHGFEMHDKKITITFASNTYUAGSNKAYIRL--NNQLMPHFVQY | 408 |
| Dy | 430 | SHEVKAGVYVAHGRCVTFASRYNCDGMKASTIHLQSSDLNRPQHR | 480 |
| RESULT 10 | | | |
| ID | ABG70122 | | |
| AC | ABG70122 | standard; protein; 494 AA. | |
| XX | ABG70122; | | |
| XX | 21-OCT-2002 | (first entry) | |
| Dy | | | |
| XX | | | |
| DE | | Human prey protein for Shigella ospC1 #40. | |
| XX | | | |
| KW | | Prey protein; ospB; ospD; ipaB; ipaC; ipaH; ospG; ospC1; Shigella; | |
| KW | | shigellosis; bacillary dysentery; antibacterial; yeast two-hybrid system; | |
| XX | | protein-protein interaction; SID; selected interacting domain; human. | |
| OS | | Homo sapiens. | |
| XX | | | |
| XX | WO200257303-A2. | | |
| XX | | | |
| PN | 25-JUL-2002. | | |
| PD | | | |
| PE | 11-JAN-2002; 2002MO-EP000777. | | |
| XX | | | |
| PR | 12-JAN-2001; 2001US-0261130P. | | |
| XX | | | |
| PA | (HYBR-) HYBRIGENICS. | | |
| XX | | | |
| PI | Legrain P; | | |
| XX | | | |
| XX | WPI; 2002-599706/64. | | |
| DR | N-PSDB; ABSS1515. | | |
| XX | | | |
| PT | | New complex of protein-protein interactions between a bait Shigella | |
| FT | | flexneri polypeptide and a prey mammalian or human placenta polypeptide | |
| PT | | for treating or preventing bacillary dysentery in a mammal or human. | |
| XX | | | |
| PS | Claim 7; Page 90-91; 162pp; English. | | |
| XX | | | |
| CC | | The invention relates to a complex of protein-protein interactions | |
| CC | | between a Shigella flexneri polypeptide (e.g. ospB, ospD, ipaB, ipaC, | |
| CC | | ipaH, 8, ospG and ospC1) and a mammalian polypeptide defined in the | |
| CC | | specification. The complexes are formed using the yeast two-hybrid | |
| CC | | system. Also included are (1) a recombinant host cell expressing the | |
| CC | | interactions between the Shigella flexneri polypeptide and a mammalian | |
| CC | | polypeptide defined in the specification; (2) selecting a modulating | |
| CC | | compound that inhibits or activates the protein-protein interactions; | |
| CC | | a modulating compound obtained from the method of (2); (4) a SID | |
| CC | | (selected interacting domain) polypeptide or its fragment or variant | |
| CC | | comprising the human polypeptides appearing as ABG70042-ABG70242; (5) a | |
| CC | | SID polynucleotide or its fragment or variant comprising encoding the | |
| CC | | above polypeptides a vector comprising (5); (6) a recombinant host cell | |
| CC | | containing the vector; and (10) a protein chip comprising Shigella | |
| CC | | flexneri polypeptide and a mammalian polypeptide defined in the | |

| CC | specification. A pharmaceutical composition comprising the compound, |
|-----------------------|--|
| CC | polymeride or polynucleotide is useful for treating or preventing |
| CC | shigellosis (bacterial dysentery) in a human or mammal. The present |
| CC | sequence represents a human prey protein isolated by the yeast two-hybrid |
| CC | assay, forming a complex of the invention with a shigella protein |
| XX | Sequence 494 AA: |
| Query Match | 16.6%; Score 570; DB 5; Length 494; |
| Best Local Similarity | 33.7%; Pred. No. 2,36-44; |
| Matches 143; | Conservative % 71; Mismatches 142; Indels 68; Gaps 16; |
| Qy | 11 PIGYRRRHQARRRMRRCNMQIQNLLEYASRQDALEYKFPNDLIKMPDAAGKXQY 70 |
| Db | 90 YIKGYRRRAASNMALGR-----PR-----AALRD-----YETVVKKPPDKDAKMKYQ 132 |
| Qy | 71 -----GSAHY-SYLDKXD-----DLVEBFG--DIVNAKIEIPIRKNHIDL 107 |
| Db | 133 ECKPIYKQKAPERAIADDEHKRSVYDSLIDBSMTIEDYISGPKLEDGKVTISFMKE----- 188 |
| Qy | 108 LIDVFRKKRGRNLHPKTYVALILREAAKSLKQLPNISPVSTVASQOVTVCGLAGKLDL 167 |
| Db | 189 LMQYVKQKQK--KLHKKCAVQILVQVKVLSKLSLTVETLKEIKETIKYCGDTHGQFYDL 246 |
| Qy | 168 VILHKNGLPSSSNPIYFVNGDFVDRGKRGLEYLLLSLYLAFENAVFLIRGNHEDSVNNA 227 |
| Db | 247 NIFELNGLPSEHTNPIYIFNGDFVDRGSSVEYILLTFPGKLLYPRPHILIRGNHETDNMQ 306 |
| Qy | 228 RYGFIRREVSERKPNHNRILAFIDEVYRWLPISGSVLSNRVLIVNGGF--SDSTSLDLTKS 285 |
| Db | 307 IYGBEGEKKATY--NQMTELSEVEEMLPALACINGKVLIMHGFLPSEGGVTLDIRK 363 |
| Qy | 286 IDRKYVSYILRPPLTDGSPDLDTKEWQDI PDIMSDPQATWGCVPNTLKGAGVWFGPDVTD 345 |
| Db | 364 IERRR-----QPR--DGGPM-----CDLWMDPQPRQNG-RSISRGVSCQFGPDVTK 407 |
| Qy | 346 NFIQRHRLSYIRSHCEKPNGEHMDNKKITITTSASNTYAIGSKKAYIRL--NNQLMPH 404 |
| Db | 408 AFLERNMIDYIIRSHVEYVABGYEVAHGRCVTVPSAPRYCDQMGNKASYIHLQGSDLRPQ 467 |
| Qy | 405 FVOY 408 |
| Db | 468 FHQF 471 |
| RESULT 11 | |
| AAE09722 | |
| ID | AAE09722 standard; protein; 499 AA. |
| XX | AAE09722; |
| XX | AC |
| XX | DT 29-NOV-2001 (first entry) |
| XX | DE Novel cell cycle protein, protein phosphatase type 5 (PP5). |
| XX | KW Cell cycle protein; DNA repair; protein phosphatase type 5; PP5; |
| XX | KW gene therapy; cancer; vaccine; cytostatic; vulnerary; antiinflammatory. |
| OS | Unidentified. |
| XX | XX |
| FH | Key Location/Qualifiers |
| FT | Misc-difference 1..7 |
| FT | /notes="Encoded by CGAG; This translation exception |
| FT | occurs while decoding with AAD16789" |
| XX | XX |
| XX | FN WO200164913-A2. |
| XX | PD 07-SEP-2001. |
| XX | PD 02-MAR-2001; 2001WO-US006849. |
| XX | PF 02-MAR-2000; 2000UG-00517779. |
| XX | XX |


```

QY 71 -----GSAHV-SVLDKX-----DLVEBFG--DIVNAKTELPFRKHIDL 107
DB 138 ECKKIYKQAKAPERAIADGDEKRSVVDLIESMTIEDYSGPKLEDEKVTISFMKE---- 193
QY 108 LIDVFRKKGNRLHPKYVALILREAAKSLKQLEPNISPVSTAVSQOVTVCGLDHLGKLDL 167
DB 194 LMQWYDQK--KLHRCACAYQILVQVKEVLSKSTLVETTLKEBTKITVCGDTHGQFYDL 251
QY 168 VILHKNGLPSSSNPYVNGDFVDRGKGLVLLLSLYAFPAVNAVFLNKGHEDSVMA 227
DB 252 NIELNGDLSEETVPYIFNGDFVDRGSPSVVEVILTLFGRKLYPDHFLILGNHETDMNQ 311
QY 228 RYGFIREVESKYPRNHRKILAFIDEVYRMPLIGSVLSRVLIVHGGF--SDSTSLDLIKS 285
DB 312 IYFEGEVKAKYT---AQWYELFSEVFEWLPLOCTINGKVLNMGGLFSEGVTLDIRK 368
QY 286 IDRGKYVSLRPPLTDGEPDLKTEWQOIFDIMSDPOATWGCVPNTLRGAGVFGPDVT 345
DB 369 IERNR-----QPF--DSGPM-----CDLMSDPQPONG-RSISKRGVSCQFGPDVT 412
QY 346 NPLQRRHLSTVIRSHCKPENGHEFMHDKITITFSASNTYAIGSNKGAAYIRL-NNOLMPH 404
DB 413 AFLERNNLDIYIIRSHVKAEGYEVANGRCVTFAPNYCDQGNKASYIHLQSSDLRPQ 472
QY 405 FVOY 408
DB 473 FHQF 476

RESULT 13
ID AEM82127 standard; protein; 499 AA.
XX AEM82127;
XX 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) polypeptide PRO25881, SEQ:5498.
XX
KM Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KM colorectal cancer; cell proliferative disorder; breast cancer;
KM colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KM central nervous system cancer; bladder cancer; pancreatic cancer;
KM cervical cancer; melanoma; leukaemia; hybridisation probe;
KM chromosome identification; chromosome mapping; gene mapping;
KM gene therapy; cytostatic.
XX
XX Homo sapiens.
XX OS
XX WO2004030615-A2.
XX PN
XX 15-APR-2004.
XX PD
XX 29-SEP-2003; 2003WO-US028547.
XX PF
XX 02-OCT-2002; 2002US-0414971P.
XX PR
XX (GETH ) GENENTECH INC.
XX XX
XX Wu TD, Zhang Z, Zhou Y;
XX PI
XX WPI, 2004-347921/32.
XX DR
XX N-PSDB; ACN40614.
XX DR
XX
XX New tumor-associated antigenic target polypeptides and nucleic acids,
XX PT useful in preparing a medicament for treating or detecting a
XX PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
XX PT prostate cancer or tumor.
XX
XX Claim 12; SEQ ID NO 5498; 7273bp; English.
XX PS
XX
XX The invention relates to human tumour-associated antigenic target (TAT)
XX polypeptides, and their related nucleic acids. The TAT polypeptides are

```

```

CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
XX represents a TAT polypeptide of the invention
XX
SQ Sequence 499 AA;
XX
Query Match 16.6%; Score 570; DB 8; Length 499;
Best Local Similarity 33.7%; Pred. No. 2.4e-44;
Matches 143; Conservative 71; Mismatches 142; Indels 68; Gaps 16;
QY 11 FIOKTRRHQAREMQRRCWQIFQNLRYASEDDQALEYFNDLIKMPQAARKNQY 70
DB 95 YIKGYRRRAASNNALQK-----FR-----AALRD-----YETVVKVPHDXDAKXQ 137
QY 71 -----GSAHV-SVLDKX-----DLVEBFG--DIVNAKTELPFRKHIDL 107
DB 138 ECKKIYKQAKAPERAIADGDEKRSVVDLIESMTIEDYSGPKLEDEKVTISFMKE---- 193
QY 108 LIDVFRKKGNRLHPKYVALILREAAKSLKQLEPNISPVSTAVSQOVTVCGLDHLGKLDL 167
DB 194 LMQWYDQK--KLHRCACAYQILVQVKEVLSKSTLVETTLKEBTKITVCGDTHGQFYDL 251
QY 168 VILHKNGLPSSSNPYVNGDFVDRGKGLVLLLSLYAFPAVNAVFLNKGHEDSVMA 227
DB 252 NIELNGDLSEETVPYIFNGDFVDRGSPSVVEVILTLFGRKLYPDHFLILGNHETDMNQ 311
QY 228 RYGFIREVESKYPRNHRKILAFIDEVYRMPLIGSVLSRVLIVHGGF--SDSTSLDLIKS 285
DB 312 IYFEGEVKAKYT---AQWYELFSEVFEWLPLOCTINGKVLNMGGLFSEGVTLDIRK 368
QY 369 IERNR-----QPF--DSGPM-----CDLMSDPQPONG-RSISKRGVSCQFGPDVT 412
QY 286 IDRGKYVSLRPPLTDGEPDLKTEWQOIFDIMSDPOATWGCVPNTLRGAGVFGPDVT 345
DB 413 AFLERNNLDIYIIRSHVKAEGYEVANGRCVTFAPNYCDQGNKASYIHLQSSDLRPQ 472
QY 405 FVOY 408
DB 473 FHQF 476

RESULT 14
ID ADE64101 standard; protein; 499 AA.
XX ADE64101;
XX AC
XX ADE64101;
XX XX
XX 29-JAN-2004 (first entry)
XX DT
XX Rat Protein P53042, SEQ ID NO 10049.
XX DE
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX KM chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
XX OS

```


QY 84 LVEFPGDIYNAKTELPIRKHIDLIDVPRKKGNNLRHPKYALLIREAKSKLQNPIS 143
 DB 167 -----GDVYTL-----DFAKMDLDFKNQK--NLHKRYAYQIVLQTRMLRALPSLV 211
 QY 144 PVSTAVSQOVTVCGLDHLGKLDLVLVLRKNGLPSSSNPYVFGDFYDRGKRGLEVLILL 203
 DB 212 DIVPBGKHFVTCGDVHGQFYDLNLIFELNGLPSEBPNPYLFNDPFVDRGSPSLSEVILTLF 274
 QY 204 SLTLAPNNAVFLNRGNHEDSVNNARYGFIREVESKYPNNHKRIIAFTIDEYTRMLPLGASVL 263
 DB 272 AFKMCPSAIIHLARGNHESKSMNKIYGFEGEVRSKLSIIFVELFA---EVFCCPLPAHVI 328
 QY 264 NSRVLLVHGG-PS-DSTSLDLKSIDRGKVSILRPPLTLDGEPDITKEMQIIFDIMWSDP 321
 DB 329 NERKVFVHRGLPSVDGDKSLDIRAIDR-----FCEPP-----EGGLMCELIWSDP 373
 QY 322 QATMGCVPTLNRGAGVWFGPDVTDNFLQRRHLSYVIRSHCKPNGHEFMNDKNTITIFSA 361
 DB 374 QPQGRGPRK-KGVGLSFGGDVTKRFLQENNLDLVRSHEVXDGBGIEIHDGKLITVFA 432
 QY 382 SNYYAIGSNKGAIVIRLN-NQLMPHFVQY 408
 DB 433 PNYCDQGNKGAFIRFEAPDMKENIVTF 460

RESULT 16
 AAE16236
 ID AAE16236 standard; protein; 485 AA.
 AC AAE16236;
 DT 09-APR-2002 (first entry)
 XX
 DE Tomato Interactor S-25 (TPR-phosphatase) protein.
 XX
 KM Tomato; signalling molecule; plant pathogen resistance; resistance gene;
 KM avirulence gene; signal transduction pathway; interactor S-25 protein;
 KM defence mechanism; PPS; TPR-phosphatase; tetratricopeptide.
 XX
 OS Lycopersicon esculentum.
 XX
 PN EPI156116-A1.
 PD 21-NOV-2001.
 XX
 PF 19-MAY-2000; 2000EP-00401402.
 XX
 PR 19-MAY-2000; 2000EP-00401402.
 XX
 PA (KEYG-) KEYGENE NV.
 XX
 PI Harling MA, Vossen JH, Simons AFM, Cornelissen BJC,
 DR WPI; 2002-084322/12.
 DR N-PSDB; AAD26447.
 XX
 PT Nucleotide sequences derived from plant genome and encoding polypeptides
 PT which interact with the product of a resistance gene are involved in
 PT signal transduction and are useful to activate plant defense against a
 PT pathogen.
 PT
 PS Disclosure; Fig 5; 66pp; English.
 XX
 CC The patent discloses nucleotide sequences derived from a plant genome,
 CC encoding polypeptides which can be activated as signalling molecules in
 CC the signal transduction pathway of plant pathogen resistance, which
 CC results from interaction between polypeptide products of a resistance
 CC gene and elicitors encoded by an avirulence gene in the pathogen and the
 CC protein interacts with the product of the resistance gene. The sequences
 CC of the invention are used to activate defence mechanisms against plant
 CC pathogens in plants. The present sequence is tomato interactor S-25 (PPS,
 CC tetratricopeptide (TPP)-phosphatase) protein

XX SQ Sequence 485 AA;
 Query Match 16.3%; Score 562.5; DB 5; Length 485;
 Best Local Similarity 35.8%; Pred. No. 1.2e-43;
 Matches 139; Conservative 55; Mismatches 125; Indels 69; Gaps 12;
 QY 24 EMORRCMQIIFQNLRYASEODAEIYKFRNDLIKHPQAAGRKQYOGSAHVSVLDDKDD 83
 DB 142 ESQRR---SVASIDYRS-----VEVEPQYAGARIE----- 169
 QY 84 LVEFPGDIYNAKTELPIRKHIDLIDVPRKKGNNLRHPKYALLIREAKSKLQNPIS 143
 DB 170 -----GDVYTL-----DFAKMDLDFKNQK--NLHKRYAYQIVLQTRMLRALPSLV 214
 QY 144 PVSTAVSQOVTVCGLDHLGKLDLVLVLRKNGLPSSSNPYVFGDFYDRGKRGLEVLILL 203
 DB 215 DIVPBGKHFVTCGDVHGQFYDLNLIFELNGLPSEBPNPYLFNDPFVDRGSPSLSEVILTLF 274
 QY 204 SLTLAPNNAVFLNRGNHEDSVNNARYGFIREVESKYPNNHKRIIAFTIDEYTRMLPLGASVL 263
 DB 275 AFKMCPSAIIHLARGNHESKSMNKIYGFEGEVRSKLSIIFVELFA---EVFCCPLPAHVI 331
 QY 264 NSRVLLVHGG-PS-DSTSLDLKSIDRGKVSILRPPLTLDGEPDITKEMQIIFDIMWSDP 321
 DB 332 NERKVFVHRGLPSVDGDKSLDIRAIDR-----FCEPP-----EGGLMCELIWSDP 376
 QY 322 QATMGCVPTLNRGAGVWFGPDVTDNFLQRRHLSYVIRSHCKPNGHEFMNDKNTITIFSA 361
 DB 377 QPQGRGPRK-KGVGLSFGGDVTKRFLQENNLDLVRSHEVXDGBGIEIHDGKLITVFA 435
 QY 382 SNYYAIGSNKGAIVIRLN-NQLMPHFVQY 408
 DB 436 PNYCDQGNKGAFIRFEAPDMKENIVTF 463

RESULT 17
 AAE09733
 ID AAE09733 standard; protein; 499 AA.
 AC AAE09733;
 DT 29-NOV-2001 (first entry)
 XX
 DE Protein phosphatase type 5 (PP5) variant, N303A.
 XX
 KM Cell cycle protein; DNA repair; protein phosphatase type 5; PP5;
 KM gene therapy; cancer; vaccine; cytostatic; vulnerary; antiinflammatory;
 KM mutant; mutelin; variant.
 XX
 OS Unidentified.
 XX
 PN Key Location/Qualifiers
 FT Misc-difference 303 /note="Wild-type Asn is substituted with Ala"
 FT
 FT W0200164913-A2.
 PN
 PD 07-SEP-2001.
 XX
 PF 02-MAR-2001; 2001WO-US006849.
 XX
 PR 02-MAR-2000; 2000US-00517779.
 XX
 PA (RIGR-) RIGEL PHARM INC.
 XX
 PI Xu X, Luo Y;
 DR WPI; 2001-557775/62.
 XX
 PT Cell cycle protein having phosphatase activity for generating agonist and
 PT antagonist e.g. antibodies of the polypeptides, both of which are useful
 PT in gene therapy and in screen assays.

XX Claim 16; Page; 75pp; English.
 PS The patent discloses cell cycle proteins and nucleic acids encoding them.
 CC The cell cycle proteins are capable of promoting cell cycle progression
 CC or cell cycle arrest. They are capable of promoting or inhibiting DNA
 CC repair. The invention also provides methods for screening a bioactive
 CC agent capable of modulating the cell cycle or interfering with the
 CC binding of a cell cycle protein, protein phosphatase type 5 (PP5). The
 CC method involves combining a cell cycle protein and a candidate bioactive
 CC agent and determining the effect on the cell in the presence or absence
 CC of the candidate agent. Cell cycle proteins bind to Rad9 proteins. They
 CC are useful for generating an agonist and antagonist of the cell cycle
 CC proteins, both of which are useful in gene therapy and in screen assays.
 CC They are useful in the diagnosis and treatment of cancer. The proteins of
 CC the invention are used as vaccines. They are useful for generating
 CC polyclonal or monoclonal antibodies to the proteins. The present sequence
 CC is N303A variant of a novel cell cycle protein, protein phosphatase type
 CC 5 (PP5) of the invention. Note: The present sequence is not shown in the
 CC specification, but is derived from the PP5 protein shown in Fig 2
 CC (AA809722)
 CC
 XX

SO Sequence 499 AA;

Query Match 16.3%; Score 562; DB 4; Length 499;

Best Local Similarity 33.5%; Pred. No. 1.4e-43;

Matches 142; Conservative 71; Mismatches 143; Indels 68; Gaps 16;

QY 11 FLOKMYRRHQAARERMRRCNMOIFONLBYASBODAEYKFFNDLIKMPAAGRNOYO 70
 DB 95 YIKGYRRRAASNMALGK-----FR-----AALRD-----YETVVKVPHDKDAKMKYQ 137
 QY 71 -----GSAHV-SVLDLKD-----DLYEERG--DIYAKTELPFRKNHIDL 107
 DB 138 ECKNIVKOKAFERAIAGDEHKRSVVDSDISEMTIDEYSGPFLBGGKTTISFMKE----- 193
 QY 108 LIDVFRKKRGNRLHPHYVALILREBAKSKLPNTISPVSTAVSQOYTVCGDLHGKLDLL 167
 DB 194 LMOWYKDKQ--KLHRCAYQILVQVKEVLSKSTLVETTLKTEKITTVCSDTHGQFYDLL 251
 QY 168 VLIHNGGLPSSSNPYFVNDGPDVDRGKGLVLLLLSLYLAPNNAVFLRGNHDSVWNA 227
 DB 252 NIFELNGLPSESTNPYIFNDGPDVDRGFSFSEVILTLFGFLYLPHDHLRGNHEDNNMQ 311
 QY 228 RYGFIREVSKPRNHRKRLAFLIDVYRWLPGLSVLNSRVLLVHGGF--SDSTSLDLIGS 285
 DB 312 IYFGSGEVAKYKT--AQMTELPSEVFEWLPPLAOCINGKVLIMHGGLFSEBDVTLDDIRK 368
 QY 286 IDRGKYVSLRPPLTDSBPLDKTEWQOIFDIMNSDPOATWGCVPNTLRGAGVFGPDVTD 345
 DB 369 IERNR-----QPP--DSGPM-----CDLWMSDPOQNG--RSISKRGVSCQFGPDVTK 412
 QY 346 NPLQRRRLSYVIRSHCKRNGHEPMDNKITTFPSNRYAIGSNKGYTRL--NNQDMH 404
 DB 413 AFLRENNLLYILIRSHKABGVAVHAGRCVTFSAPNYCDQMGNTASYIHLQGSDLRQ 472
 QY 405 FVOY 408
 DB 473 FHQF 476

RESULT 18

AA809731

AA809731 standard; protein; 499 AA.

XX

AC AA809731;

XX 29-NOV-2001 (first entry)

XX Protein phosphatase type 5 (PP5) variant, H244A.

DB Cell cycle protein; DNA repair; protein phosphatase type 5; PP5;

XX gene therapy; cancer; vaccine; cytostatic; vulnary; antiinflammatory;

KW mutant; mutcin; variant.

XX Unidentified.

OS Key Location/Qualifiers

FT Misc-difference 244 /note="wild-type His is substituted with Ala"

XX MO200164913-A2.

XX 07-SEP-2001.

XX 02-MAR-2001; 2001WO-US006849.
 XX 02-MAR-2000; 2000US-00517779.
 XX (RIGB-) RIGEL PHARM INC.
 XX Xu X, Luo Y;
 XX MPI; 2001-557775/62.
 DR Cell cycle protein having phosphatase activity for generating agonist and
 PT antagonist e.g. antibodies of the polypeptides, both of which are useful
 PT in gene therapy and in screen assays.
 XX Claim 16; Page; 75pp; English.
 CC The patent discloses cell cycle proteins and nucleic acids encoding them.
 CC The cell cycle proteins are capable of promoting cell cycle progression
 CC or cell cycle arrest. They are capable of promoting or inhibiting DNA
 CC repair. The invention also provides methods for screening a bioactive
 CC agent capable of modulating the cell cycle or interfering with the
 CC binding of a cell cycle protein, protein phosphatase type 5 (PP5). The
 CC method involves combining a cell cycle protein and a candidate bioactive
 CC agent and determining the effect on the cell in the presence or absence
 CC of the candidate agent. Cell cycle proteins bind to Rad9 proteins. They
 CC are useful for generating an agonist and antagonist of the cell cycle
 CC proteins, both of which are useful in gene therapy and in screen assays.
 CC They are useful in the diagnosis and treatment of cancer. The proteins of
 CC the invention are used as vaccines. They are useful for generating
 CC polyclonal or monoclonal antibodies to the proteins. The present sequence
 CC is H244A variant of a novel cell cycle protein, protein phosphatase type
 CC 5 (PP5) of the invention. Note: The present sequence is not shown in the
 CC specification, but is derived from the PP5 protein shown in Fig 2
 CC (AA809722)
 CC
 XX

SO Sequence 499 AA;

Query Match 16.3%; Score 560; DB 4; Length 499;

Best Local Similarity 33.5%; Pred. No. 2.1e-43;

Matches 142; Conservative 71; Mismatches 143; Indels 68; Gaps 16;

QY 11 FLOKMYRRHQAARERMRRCNMOIFONLBYASBODAEYKFFNDLIKMPAAGRNOYO 70
 DB 95 YIKGYRRRAASNMALGK-----FR-----AALRD-----YETVVKVPHDKDAKMKYQ 137
 QY 71 -----GSAHV-SVLDLKD-----DLYEERG--DIYAKTELPFRKNHIDL 107
 DB 138 ECKNIVKOKAFERAIAGDEHKRSVVDSDISEMTIDEYSGPFLBGGKTTISFMKE----- 193
 QY 108 LIDVFRKKRGNRLHPHYVALILREBAKSKLPNTISPVSTAVSQOYTVCGDLHGKLDLL 167
 DB 194 LMOWYKDKQ--KLHRCAYQILVQVKEVLSKSTLVETTLKTEKITTVCSDTHGQFYDLL 251
 QY 168 VLIHNGGLPSSSNPYFVNDGPDVDRGKGLVLLLLSLYLAPNNAVFLRGNHDSVWNA 227
 DB 252 NIFELNGLPSESTNPYIFNDGPDVDRGFSFSEVILTLFGFLYLPHDHLRGNHEDNNMQ 311
 QY 228 RYGFIREVSKPRNHRKRLAFLIDVYRWLPGLSVLNSRVLLVHGGF--SDSTSLDLIGS 285
 DB 312 IYFGSGEVAKYKT--AQMTELPSEVFEWLPPLAOCINGKVLIMHGGLFSEBDVTLDDIRK 368


```

QY 72 SAHVSDDKD---DLVE-----BFGDVIANKIEIPRKHNDLID---- 110
DQ 190 -----DNDATKALVECEKIVQQLAPFAAIEVGDLSAAEGLDVSMADVASYDVR 242
QY 111 -----VFRKRGNRILPKYVALILREBAKSLQLEPISFVSTAVSQVTCG 157
DQ 243 EGNEMTQEFIDMIEFRRKGLIHKYVYQIIIVRNIVYNEFTMEVDIPBDVQLTVCG 302
QY 158 DLNG-----KLDLVLVHKRLPSSSNPYVNGPVDGRKGL 197
DQ 303 DTGGMFALAMPPEXWLLTNCXAKYQYDLMELFRLNFPSPDKHYTLFNGDFVDRGSWTE 362
QY 198 VLLLSLYLAFPAVAFNKGHEDSVNARVGFIREVSKY----- 239
DQ 363 IALLLVAYKRLRNGFINGNHETDMNRYVGFEGECCKAKYERASRLCAHYATSKI 422
QY 240 -----PRNHRILAFIDEVYRWLPGLSVLSRVLYVHGPF--SDSTSLDLIKS 285
DQ 423 FEHTPADTSHLPRYK-----LFSFSPSALPLATLICKFVLVHGLFSDDNVTLDIRK 478
QY 286 IDRGKVISILRPPLTGEPLDKTEWQOIFDIMSDDPQATMGCVPTNLRGAGVFGPDVT 345
DQ 479 LDRHKQ---KQPGQAG-----LMEMELWTDPPQPPGRGSPSK-RGVGMQFGPDVT 524
QY 346 NFLQHRILSYIRSHCEKPNHGFEMDNKIITIPASANYAIGSNKAYIRLANNQMLPHF 405
DQ 525 RFDCKGKGLBAILIRSHVEMDGYEBEHDKCITVFSAPKCYDMTENKAYINIGPDYKGF 584
QY 406 VQY 408
DQ 585 SQF 587

RESULT 22
ID ABG09988 standard; protein; 714 AA.
XX AC ABG09988;
XX DT 13-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #9979.
XX KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN W0200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001MO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI, 2001-639362/73.
XX DR N-PSDB; AAS74175.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 40347; 103bp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridization probes, polymerase chain

```

```

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantifying a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC patent did not appear in the invention. Note: The sequence data for this
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 714 AA;
XX
XX Query Match 14.4%; Score 496.5; DB 4; Length 714;
XX Best Local Similarity 29.2%; Pred. No. 3.9e-37;
XX Matches 140; Conservative 62; Mismatches 147; Indels 131; Gaps 15;
QY 41 SEQDOAEYKFFNDLIKHPQAGRNQYQ-----GSAHV--SVLDDKD- 82
DQ 231 SEHEAATVYL---VKKVPHDKDAKQYQCKKIYKQKAFERALNGDEHKRSVNSLDI 287
QY 83 ---DLVEBFG--DIYNAKIEIPRKHNDLIDVFRKRGNRILPKYVALILREBAKSLK 137
DQ 288 ESWTIDEVSGPQLBEGKVTISFMEK-----LMQWYDQK--KLHRCAYQILVQVKEVLS 341
QY 138 QLEPNISPVTAAGVQVTCGDLHGKLDLVLVHKNGLPSSSNPY----- 183
DQ 342 KLSLVETTLKERTKLTICGDTHGQFYDLNIFELNGLSPSTNPTYSFSAPEPSPHLPQ 401
QY 184 -----FNGDPVDRGKGLVLLLSLYLAP 209
DQ 402 PQQVCTWEBEPLRGKDVTLERRGLAMLGHQGI FNGDPVDRGSFSEVVLTLFGPRLV 461
QY 210 PNAVFLNRGNHESVNAARYGPIRVEESKYPRNHRILAFIDVYRWLPGLSVLSRVLI 269
DQ 462 PDHFIHLRGNHEDNNQIYGFEGEVKAKYT--AQMYELFSVFEMPLPLAQCIKSKVLI 518
QY 270 VHGGF--SDSTSLDLIKSIDRGKVISILRPPLTGEPLDKTEWQOIFDIMSDDPQ----- 322
DQ 519 MHGGLFSEBQVTLDDIRKTERNR-----QPP--DSALPRRMGFPILARRSGSSPQPPHLR 571
QY 323 -----ATMGCVPTNL-----RGAGVWFGPDVTNPLQ 349
DQ 572 LGPVHLHGAVMASTVTVTVVRGEGVTAHTCPYLFNGRSISRGVSCQGRPVYTAFLR 631
QY 350 RHRILSYIRSHCEKPNHGFEMDNKIITIPASANYAIGSNKAYIRL--NNQMLPHFVQY 408
DQ 632 ENNLDIYIRSHVEKAGYEVAGRCVTVFSAPNYCDQGNKASYIHLQSSDLRPQFHQF 631

RESULT 23
ID ABR53882 standard; protein; 513 AA.
XX AC ABR53882;
XX DT 20-JUN-2003 (first entry)
XX DE Protein sequence #SEQ ID 2629.
XX KM Multiprotein complex; eukaryote; drug target; diagnostic.
XX OS Saccharomyces cerevisiae.

```


| | | | |
|-----------|---|--|-----|
| Oy | 177 | SSNPVFNFGDVFDRSGKGLBVLTLTLSTYLAPNMAVFLNRGNHBSVMNARGFIREVE | 238 |
| Db | 268 | GRKATVLEFGDFDRSGWSCEVALVLYCLKTLHPNPFILNRGNHSDNMKTIQPEDECK | 327 |
| Oy | 237 | SKYPRNKKIILAFIEVRYWMLPLGYVANSRVLLVHGGF-SD-STSLDILISIDRGKYVS | 294 |
| Db | 328 | YKY----SQGIFPMNFAQSFBSLPLATILINDYLVHGGCLPBDPATILSDPFNIDR-----F | 379 |
| Oy | 235 | LRPPLTDEGPLDKTEWQOIFDIMWSDPQATMGCVENLTARGAVMFGPDVTNDFLQRRHLS | 354 |
| Db | 380 | AQPP-RDG-----AFMEILLMAADPOEANGMGPSQ-RGLGHAFGPDIDRFLRNKLR | 428 |
| Oy | 355 | VIIRSECKNGHEFMHDKIITIFASANYVALGSKGAVTRL | 397 |
| Db | 429 | KIFRSHELMGQVQEPQKGLMTVFSAFNPYCDSSQGNLGVYIHV | 471 |
| RESULT 25 | | | |
| ADSA43550 | | | |
| XX | ADSA43550 | standard; protein; 513 AA. | |
| XX | | | |
| XX | ADSA43550; | | |
| DT | 02-DEC-2004 | (first entry) | |
| XX | | | |
| DE | Bacterial polypeptide #21980. | | |
| XX | | | |
| KW | Recombinant DNA construct; transformed plant; improved plant property; | | |
| KW | cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis; | | |
| KW | pathogen tolerance; pest tolerance; plant disease resistance; | | |
| KW | cell cycle pathway modification; plant growth regulator; | | |
| KW | homologous recombination; seed oil yield; protein yield; carbohydrate; | | |
| KW | nitrogen; phosphorus; photosynthesis; lignin; galactomannan; | | |
| XX | bacterial polypeptide. | | |
| OS | Bacteria. | | |
| XX | | | |
| PN | US2003233675-A1. | | |
| XX | | | |
| PD | 18-DEC-2003. | | |
| XX | | | |
| PE | 20-FEB-2003; 2003US-00369493. | | |
| XX | | | |
| PR | 21-FEB-2002; 2002US-0360039P. | | |
| PA | (CAOY/) CAO Y. | | |
| PA | (HINK/) HINKLE G J. | | |
| PA | (SLAT/) SLATER S C. | | |
| PA | (CHEN/) CHEN X. | | |
| PA | (GOLD/) GOLDMAN B S. | | |
| PI | Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS; | | |
| XX | | | |
| DR | WPI, 2004-061375/06. | | |
| XX | | | |
| PT | New recombinant DNA construct comprising a promoter positioned to provide | | |
| PT | for expression of a polynucleotide encoding a polypeptide from a | | |
| PT | microbial source, useful for producing plants with improved properties. | | |
| XX | | | |
| XX | Claim 1; SEQ ID NO 21980; 122pp; English. | | |
| XX | | | |
| CC | The invention relates to a recombinant DNA construct comprising a | | |
| CC | promoter functional in a plant cell, where the promoter is positioned | | |
| CC | to provide for expression of a polynucleotide encoding a polypeptide from a | | |
| CC | microbial source. The invention also relates to a transformed plant | | |
| CC | comprising the recombinant DNA construct and a method of producing a | | |
| CC | transformed plant having an improved property. The plant is a crop plant | | |
| CC | such as maize or soybean. The method of producing a transformed plant | | |
| CC | having an improved property comprises transforming a plant with the | | |
| CC | recombinant DNA construct and growing the transformed plant, where the | | |
| CC | polynucleotide or polypeptide is useful for improving plant properties. | | |
| CC | The recombinant DNA construct is useful for producing plants with | | |
| CC | improved plant properties, e.g. improved cold, heat or drought tolerance, | | |

| | | |
|---------------------------|---|--|
| CC | tolerance to herbicides, extreme osmotic conditions, pathogens or pests, | |
| CC | increased resistance to plant disease, better growth rate by modification | |
| CC | of the cell cycle pathway with plant growth regulators, increased rate of | |
| CC | homologous recombination, modified seed oil or protein yield and/or | |
| CC | content, improved yield by modification of carbohydrate, nitrogen or | |
| CC | phosphorus use and/or uptake, by modification of photosynthesis or by | |
| CC | providing improved plant growth and development under at least one stress | |
| CC | condition, improved lignin production or improved galactomannan | |
| CC | production. This sequence represents a bacterial polypeptide used in the | |
| CC | scope of the invention. Note: The sequence data for this patent did not | |
| CC | form part of the printed specification but was obtained in electronic | |
| CC | format from USPTO at seqdata.uspto.gov/sequence.html . | |
| XX | | |
| SQ | Sequence 513 AA; | |
| Query Match | 13.9%; Score 479.5; DB 8; Length 513; | |
| Best Local Similarity | 32.3%; Pred. No. 9.3e-36; | |
| Matches 130; Conservative | 68; Mismatches 154; Indels 51; Gaps 13; | |
| QY | 2 DENAIRAIFIQKYYRHHQARREM---QRRNCWOI PQNLAYASBODQAEIYKFFNDLIKH 58 | |
| DB | 113 DPAATKALLTCDFRFRBERFRKAIIGAENBAKISLCQTLNLSFDPANADILANY----- 165 | |
| QY | 59 MPQAAGRNQVQGSANHSVLDKDDVBEFDIYNAKIELPIRKHHIDLIDVPRKKRGN 118 | |
| DB | 166 ---EGPDLBEP-----QLYDDKN-----AFKGAIKIKNNSQEPFISKQVNDLFLK-GK 207 | |
| QY | 119 RLHPKCYVALIRBAKASIKOLPNISPV--SPAVSQOVTVCGLDKGKLDLLVWLHKNGLP 176 | |
| DB | 208 YLPKKTVAALISHADTLFRQEPSPWELENSTPDKVKSVCDDITGQFYDVLNTPRKRGKV 267 | |
| QY | 177 SSSNPYVNGDPVDVGRKKGLVLLLLSLYLAPFNAAVFLKNGHEDSVNNAKYPIREVE 236 | |
| DB | 268 GPKHTYLPFNQDFVDRGWSCEVALLFYCLKILHPNNPFLKNGHESDMNKKIYGPBEDECK 327 | |
| QY | 237 SKYPRNKRGIILAFIDEVYRMLPLQSLVNSRPLIYHGGP-SP-STSLDLIKSIDRGKTVSI 294 | |
| DB | 328 YKY---SQRIINMPAQSPESLPLATLINNDYLVNHGGLPSPPSATLSDPFKNIDR-----F 379 | |
| QY | 295 LRPLPTDGEPLDKTKEWQOIPDIMMSDPQATWGCVPNTLRGAGWFGPVDVTDNPLQRRHLS 354 | |
| DB | 380 AQP-RD-----APMELMADPQBANGMGSPQ-RGLGHAFGPDITDRFLNNCLR 428 | |
| QY | 355 YVIRSHCEKPNHGHEFMHDKIITTFASANYAIGSNKGAIVTRL 397 | |
| DB | 429 KIFRSHELRMGCVGPBEQKGLMTYFSAPNYCDSGNLGVYIHV 471 | |
| XX | | |
| RESULT 26 | | |
| ABBS9732 | | |
| ID | ABBS9732 standard; protein, 346 AA. | |
| XX | | |
| AC | ABBS9732; | |
| XX | | |
| DT | 26-MAR-2002 (first entry) | |
| XX | | |
| DE | Drosophila melanogaster polypeptide SEQ ID NO 5988. | |
| XX | | |
| KM | Drosophilid; developmental biology; cell signalling; insecticide; | |
| XX | pharmaceutical. | |
| XX | | |
| OS | Drosophila melanogaster. | |
| XX | | |
| PN | WO200171042-A2. | |
| XX | | |
| PD | 27-SEP-2001. | |
| XX | | |
| PF | 23-MAR-2001; 2001WO-US009231. | |
| XX | | |
| PR | 23-MAR-2000; 2000US-0191637P. | |
| XX | | |
| PR | 11-JUL-2000; 2000US-00614150. | |
| XX | | |
| PA | (PERR) PE CORP NY. | |

Db 222 IGPELKLSEYQVFAVPH 238

RESULT 28
ABP98033
ID ABP98033 standard; protein; 323 AA.
XX
XX ABP98033;
AC
DT 11-AUG-2003 (first entry)
XX
XX Protein phosphatase stress-related polypeptide GmPP2A-4.
DE
XX
XX Protein phosphatase stress-related polypeptide; PPSRP; PpPP2A-1; PpPP-1;
KW BmPP2A-1; BmPP2A-2; BmPP2A-3; GmPP2A-1; GmPP2A-2; GmPP2A-3; OsPP2A-1;
KM OsPP2A-2; OsPP2A-3; plant tolerance; environmental stress; plant;
KW drought; salinity; cold; enzyme.
XX
XX Glycine max.
OS
XX
XX WO2003020914-A2.
PN
XX
XX 13-MAR-2003.
PD
XX
XX 05-SEP-2002; 2002MO-US028445.
PF
XX
XX 05-SEP-2001; 2001US-0317305P.
PR
XX
XX (BADI) BASF PLANT SCI GMBH.
PA
PI Bohmert HJ, Chen R, Ishitani M, Van Thielens N, Da Costa E SilvaO;
XX
XX WPI; 2003-30086/29.
DR
XX
XX N-PSDB; ACC43348.
PT
XX
XX New protein phosphatase stress-related polypeptide coding nucleic acid,
PT useful for modulating plant's tolerance to an environmental stress such
as drought, increased salinity and cold.
XX
XX
XX Claim 13; Page 81; 107P; English.

CC The present sequence represents a protein phosphatase stress-related
CC polypeptide (PPSRP). The specification describes PPSRP polypeptides
CC designated PpPP2A-1, PpPP-1, BmPP2A-1, BmPP2A-2, BmPP2A-3, GmPP2A-1,
CC GmPP2A-2, GmPP2A-3, OsPP2A-2 or OsPP2A-3, derived from
CC Physcomitrella patens, Brassica napus, Glycine max or Oryza sativa. PPSRP
CC polypeptides and polynucleotides are useful for modulating plant
CC tolerance to an environmental stress such as drought or increased
CC salinity and cold. They are also useful in identification and
CC localization of Physcomitrella patens, Brassica napus, Glycine max or
CC Oryza sativa and related organisms, mapping of genomes of organisms
CC related to the above species, in a evolutionary and polypeptide
CC structural studies, in determination of PPSRP regions required for
CC function, modulation of PPSRP activity, modulation of metabolism of one
CC or more cell functions and transmembrane transport of one or more
CC components
CC
XX
XX Sequence 323 AA;
SQ

Query Match 12.7%; Score 435.5; DB 6; Length 333;
Best Local Similarity 31.2%; Pred. No. 6,4e-32;
Matches 104; Conservative 66; Mismatches 124; Indels 39; Gaps 9;

QY 90 DIVNAKIELP-IRKNHIDILIDYRKKRGRIHFKYVALIRBAKSLKQLPINSPYSTA 148
DB 8 DILNRLEAVPTLPKQYQLSBSRIQ-----LCVVSRI--FLQQPNLIREAP 54
QY 149 VSOQVTVCGDLGKLDLVLVHKGGLPSSSNPVFVGSDVDRGKGLVLLLLSLVLA 208
DB 55 IK-----ICGVHGGYSDLLRLPEYGLPPEAN-YLFGDYVDREKQSLFTICLLAVKIK 109
QY 209 PPAAVPLNRGNHDSVNAKYGPIREVSCKIPRHKRIILAFIDEVYRWLPLGVSATNSRVL 268

Db 110 YPENPFLKGNHCEASINRIYGYDECKRRP---NWLMTFTTDCRCFLVAAALVDEKIL 166
QY 269 IVHGFS-DTSLDLKSIDRGKYSILRPPLTDGEPDTEBMOQIFDIMSDPOATWGC 327
DB 167 CMHGGLSPDLNNDQIRNLCR-----PTDVPDVGILCDLMSDPSKXVCG 211
QY 328 VPNTLAGAGVFEPPDYTNFLQHRRLSYVIRSHCEKNGHEFMHDKIITIFSAASYAI 387
DB 212 WGNDRGVSYTFGADRVQSPLQGHDDLVCRAHQVEDGDEYFPMNQLVITFSAPNYCGR 271
QY 388 GSNKAYIRLNLMPHFVQYISAASOTKLSF 420
DB 272 FDNAGMMSVDETLMCSF-QILKPADKAKALNP 303

RESULT 29
ADX87721
ID ADX87721 standard; protein; 339 AA.
XX
XX ADX87721;
AC
XX
XX 21-APR-2005 (first entry)
DT
XX
XX
XX Plant full length insert polypeptide seqid 50385.
DE
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
XX
XX Unidentified.
OS
XX
XX US2004034888-A1.
PN
XX
XX 19-FEB-2004.
PD
XX
XX 28-APR-2003; 2003US-00425114.
PF
XX
XX 06-MAY-1999; 99US-00304517.
PR
XX
XX 05-NOV-2001; 2001US-00985678.
XX
XX
XX (LIU/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAS/) TABASKA J E.
PA (CAOV/) CAO Y.
PI
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
XX WPI; 2004-180133/17.
DR
XX
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX
XX Claim 1; SEQ ID NO 50385; 15P; English.

CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring

CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 339 AA;
Query Match 12.6%; Score 434.5; DB 8; Length 339;
Best Local Similarity 30.1%; Pred. No. 8.7e-32;
Matches 106; Conservative 72; Mismatches 135; Indels 39; Gaps 9;
QY 56 IKHMPQ-AAQRKQYQSAHVSVDKDLVEBFQDIYNAKIELPIRKHIDLLIDVPRK 114
DB 1 LRIHPPSPDGRK---KGSGKMMWTR-----APMGPMGAADVEMVR-----LVGGGKG 46
QY 115 KGGNRLHPKYVALILBEAKSLKQLPNISPVSTAVSQYTVGGDLKGLDILLVLAHKG 174
DB 47 GRVQLSBAIRQLCVBGRKVLISQPNL---LRHAPVKICGDHGFVDLRLRFDLQ 102
QY 175 LPSSSNPYVFGDFVDGRKGLFVLLLLSLYLAFPAFVLRKGNHDSVMAARYGFI 234
DB 103 YPPAGT-YVFLGDYVDGRKQSLFTICLLAYKIRYENFIKRGNHEDAKIRYVGFY 161
QY 235 VESKTYRNHRIAFIDEVYRWLPDGSVLSRVLIYHGFS-DTSLDLIKSIDRGKYS 293
DB 162 CKRRF---NVRLMKIPSDCFNCLPIALIDDKILCHNGGISPETSLDOIKER----- 213
QY 294 ILRPPLDSEPTDKTEMQQIFDIWMSDPQATWGCVENTLRGAGVWPGPYTNELQRHL 353
DB 214 -----PAIPDYGLCDLMSDPSPDGGMGESDRGVSCTGAKLVFEKNDL 263
QY 354 SYVIRSECKPNGHSEPMHDKITITFSASNYVAIGSKGAYIRLNQLMPHF 405
DB 264 DLICRHHQVDEQYEFPAQRRLVITFSAPYCGEFPNVGALLSIDSLMCSF 315
RESULT 30
AAG30062
ID AAG30062 standard; protein, 318 AA.
XX
AC AAG30062;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35873.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0133048P.
PR 30-APR-1999; 99US-0133407P.
PR 04-MAY-1999; 99US-0133484P.
PR 05-MAY-1999; 99US-0133485P.
PR 06-MAY-1999; 99US-0133486P.
PR 06-MAY-1999; 99US-0133487P.
PR 07-MAY-1999; 99US-0133863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140629P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142052P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-015087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145227P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 28-JUL-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.

PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161929P.
 PR 28-OCT-1999; 99US-0161933P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 12.6%; Score 434; DB 3; Length 318;
 Best Local Similarity 29.6%; Pred. No. 8,7e-32;
 Matches 106; Conservative 68; Mismatches 120; Indels 64; Gaps 11;

QY 59 MPQARKNQYQSAHVSVDKDDVVEEGDVIANKIEPIKKNHIDLDIVPRKR-- 116
 DB 1 MAEKPAQDEQKQKAMEBPAVD-----DIIRLVR-----FRNTRPG 36
 QY 117 -GNRLHPKYVALILREAKSLKQLPNIS-----PVSTAVSQQVTCGDHKGDLIDLIV 168
 DB 37 SGKQVH-----LSGG--EIRQLCAVSKRIPIQGNLLELEAPIKICGDIHQYSDLR 87
 QY 169 VLRKNGLPSSSNPYVNGDFVDRGKRGLEYLLLSLYLAFAVPAVPLNRGNHDSVMNAR 228
 DB 88 LPEYGGFPPEAN-YLFGVDVDRGKQSLFTICLLAVKIKYPNPFILRGNHBSASINRI 146
 QY 229 YGPIRVESEKYPNNHRIILAFIDEVYRWLPLGSVANSRVLIYVGRS-DSTSLDKSID 287
 DB 147 YGPIYDECKRRF--NRLWKIFITDCNCLPVALIDRLICMGGISPELKSIDQIRNIA 203
 QY 288 RGYKVSILRPPLTDGSPLDKTEWQOIFDIMWSDPOATMGCPPTLKGAGVWPGPDVTDNF 347
 DB 204 R-----PMDPESGLVCDLMSDPGVDGKGMND-RGVSTTFGADKYAEF 247
 QY 348 LQRRRLSYVIRSHCEKNGHEFMHDKIITTFASANYTALGSKGAYIRLNQJLMPHF 405
 DB 248 LRRKMDLICRAHQVVEDGYEFPAERQLVTVFAPVNYCGBFNAGAMMSIDBSLMCSF 305

RESULT 31
 ADN72465
 ID ADN72465 standard; protein; 318 AA.
 XX
 AC ADN72465;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Thale cress protein upregulated in E2Fa/Dpa expressing plants Segid 360.
 XX
 KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;
 KM animal feed product; thale cress; cell wall biosynthesis;
 KW nitrogen metabolism; carbon metabolism.
 OS Arabidopsis thaliana.
 OS
 PN WO2004035798-A2.
 XX
 PD 29-APR-2004.

[illegible]

ID ADY09284 standard; protein; 380 AA.
 XX
 XX AC ADY09284;
 XX
 DT 21-APR-2005 (first entry)
 DE Plant full length insert polypeptide seqid 65099.
 XX
 XX plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomanan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content.
 XX
 OS unidentified.
 XX
 PN US200403488-A1.
 XX
 PD 19-FEB-2004.
 XX
 PF 28-APR-2003; 2003US-00425114.
 XX
 PR 06-MAY-1999; 99US-00304517.
 PR 05-NOV-2001; 2001US-00385678.
 XX
 PA (LITU/) LITU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCRE S E.
 PA (TABAA/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovacic DK, Screen SE, Tabaska JE, Cao Y;
 DR WPI; 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 PS Claim 1; SEQ ID NO 65099; 15pp; English.
 XX
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp://seqdata.uspto.gov/sequence.html?pcid:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomanan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.
 XX
 SQ Sequence 380 AA;
 XX
 Query Match 12.5%; Score 431.5; DB 8; Length 380;
 Best Local Similarity 30.5%; Pred. No. 2e-31;
 Matches 102; Conservative 66; Mismatches 125; Indels 41; Gaps 8

Db 62 DIINRLLEV-----RSRPGKQVQLSESRIRHLCAASREIFLQGPVLLLELA 107

Qy 148 AVSQQVTCGDLHGKLDLVLVLRKNGLPSSSNPYVNGDFVDRGKGLVLLLSLYL 207

Db 108 PIR-----ICGVHGHQVSDLLRLFPYGGI.PPEAN-YLFLGIDYVDRGKSLFTICLLAVKI 162

Qy 208 AFPAVFLNRGNHSDVMNANRGFIRVESKYPNNHRILAFIDVYRMVPLGSVLSRV 267

Db 163 KTFENFLRGNHSCASINRIYGVYDECKRRF--NRLMTPTTECRNCLFVALIDEKI 219

Qy 268 LIVHGFS-DSTSLDLIKSIDRGKYVILRPPLTDGSPLDKTEWQQLFDIMWSPDQATMG 326

Db 220 LCHHGGLSPDILNDQIRNLQR-----PTDVPDVGILCDLIMSPSPSEVQ 264

Qy 327 CVPTLRGAGVFGPDVTDNLFQRRLSYVIRSHCKPNHGFPMNDKIIITIFASNYTA 386

Db 265 GWMGNDRGVSTYFGADKVEFLQKHGDDLICRAHQVVEDGVFPANRQLVITFSAFNYCG 324

Qy 387 GSNKGAIVIRLNQMLMPHFVQYISASQTKLSF 420

Db 325 EPDNGAMMSVDETLMCSF-QILKPADKKAALNF 357

RESULT 33

ADT59840

ID ADT59840 standard; protein; 325 AA.

XX ADT59840;

AC

XX

DT 13-JAN-2005 (first entry)

XX

DE Plant polypeptide, SEQ ID 9917.

XX

KM Plant; transgenic; cold tolerance; growth rate; drought tolerance;

KM disease resistance; galactomanan production; plant growth regulator;

KM heat tolerance; herbicide tolerance; lignin production;

KM extreme osmotic condition tolerance; pathogen resistance;

KM pest resistance; yield improvement; seed oil yield; seed protein yield.

XX

OS Viaridiplantae.

XX

PN US2004216190-A1.

XX

PD 28-OCT-2004.

XX

PF 18-DEC-2003; 2003US-00739930.

XX

PR 28-APR-2003; 2003US-00424599.

XX

PR 28-APR-2003; 2003US-00425115.

XX

PA (KOVA/) KOVALIC D K.

XX

PI Kovalic DK;

XX

PI WPI; 2004-757369/74.

XX

DR

XX

PT New recombinant DNA constructs useful in the field of biochemistry and

PT genetics, and in particular for producing transgenic plants with improved

PT biological characteristics.

XX

PS Claim 2; SEQ ID NO 9917; 14pp; English.

XX

XX The invention relates a recombinant DNA construct comprising a

CC polynucleotide having any of 5544 nucleotide sequences (CDNAs SEQ ID NO:

CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences

CC (SEQ ID NO: 5545-11088). The CDNAs and proteins are from corn, soybean,

CC Arabidopsis, wheat and rape but the specification does not indicate which

CC sequences is derived from which organism. Also included is a method of

CC producing a plant having an improved property, comprising transforming a

CC plant with a recombinant DNA construct comprising a promoter region

CC functional in a plant cell operably joined to a polynucleotide encoding a

CC polypeptide associated with the property, and growing the transformed

CC plant. The property is selected from improving plant cold tolerance, for

CC manipulating growth rate in plant cells by modification of the cell cycle

CC pathway, for improving plant drought tolerance, for providing increased

CC resistance to plant disease, for galactomanan production, for production

CC of plant growth regulators, for improving plant heat tolerance, for

CC improving plant tolerance to herbicides, for increasing the rate of

CC homologous recombination in plants, for lignin production, for improving

CC plant tolerance to extreme osmotic conditions, for improving plant

CC tolerance to pathogens or pests, for yield improvement by modification of

CC photosynthesis, for modifying seed oil yield and/or content, for

CC modifying seed protein yield and/or content, for yield improvement by

CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake

CC and for yield improvement by providing improved plant growth and

CC development under at least one stress condition. The polynucleotide may

CC also encode a plant transcription factor. The methods and compositions of

CC the present invention are useful in the field of biochemistry and

CC genetics, in particular for producing transgenic plants with improved

CC biological characteristics such as increased yield, improved nitrogen

CC flow, increasing plant tolerance to cold or heat, improving plant

CC tolerance to extreme osmotic and drought conditions, and improving plant

CC tolerance to plant pests or pathogens. They can also be used in physical

CC arrays of molecules, plant breeding markers, computer-based storage and

CC analysis systems. The present sequence is one of the 5544 plant protein

CC sequences of the invention. Note: The sequence data for this patent did

CC not form part of the printed specification, but was obtained in

CC electronic format directly from USPTO at

CC seqdata.uspto.gov/sequence.html?docid=2040216190.

XX

SQ Sequence 325 AA;

XX

Query Match 12.5%; Score 430.5; DB 8; Length 325;

Best Local Similarity 31.1%; Pred. No. 1,9e-31;

Matches 99; Conservative 69; Mismatches 122; Indels 29; Gaps 6;

Qy 89 GDIWNAKIELPIRKNIHIDLIDVFRKKGRNRLHPKVVALIRBAASLKLQPLNISPVSTA 148

Db 9 GAMEGALNDVYVR-----LVBCGRGRQVQLSEARIRQLCDVAKRVFLSQPMU-----LR 59

Qy 149 VSOQVTCGDLHGKLDLVLVLRKNGLPSSSNPYVNGDFVDRGKGLVLLLSLYLA 208

Db 60 IQAPVVICGDIHQFVDLRLFLDGLGVPPST-YVFLGIDYVDRGKSLFTICLLAVKIR 118

Qy 209 FPAVFLNRGNHSDVMNANRGFIRVESKYPNNHRILAFIDVYRMVPLGSVLSRV 268

Db 119 YPDKVFLLRGNHSDAIVNRVGYDECKRRF--NRLMTPTTECRNCLFVALIDEKI 175

Qy 269 IVHGFS-DSTSLDLIKSIDRGKYVILRPPLTDGSPLDKTEWQQLFDIMWSPDQATMG 327

Db 176 CHHGGLSPDILNDQIRNLQR-----PAPIDVVGILCDLIMSPSPSEVQ 260

Qy 328 VPPTLRGAGVFGPDVTDNLFQRRLSYVIRSHCKPNHGFPMNDKIIITIFASNYTA 387

Db 221 WGSDDRGVSTYFGADKVEFLQKHGDDLICRAHQVVEDGVFPANRQLVITFSAFNYCG 280

Qy 388 GSNKGAIVIRLNQMLMPHF 405

Db 281 PDNVGALSTIDETLMCSF 298

RESULT 34

ADY08660

ID ADY08660 standard; protein; 334 AA.

XX

AC ADY08660;

XX

DT 21-APR-2005 (first entry)

XX

DE Plant full length insert polypeptide seqid 64475.

XX

KM plant protectant; plant growth regulant; gene therapy; plant;

KM recombinant DNA construct; physical array; plant breeding marker;

KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;

KM extreme osmotic condition; pathogen tolerance; pest tolerance;

KM growth rate; cell cycle pathway; disease resistance;

KM galactomannan production; lignin production; plant growth regulator;
 KM yield; plant growth; plant development; seed oil; protein yield;
 KM protein content.
 OS Unidentified.
 XX
 XX
 PN US2004034888-A1.
 XX
 XX
 PD 19-FEB-2004.
 XX
 PF 28-APR-2003; 2003US-00425114.
 XX
 XX 06-MAY-1999; 99US-00304517.
 PR 05-NOV-2001; 2001US-00985678.
 XX
 PA (LIU/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABA/) TABASKA J E.
 PA (CAOV/) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y,
 DR MPI, 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 64475; 15pp; English.
 XX
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.
 CC
 XX
 XX Sequence 334 AA:
 SQ
 Query Match 12.5%; Score 430.5; DB 8; Length 334;
 Best Local Similarity 30.6%; Pred. No. 2e-31;
 Matches 113; Conservative 67; Mismatches 134; Indels 55; Gaps 13;
 QY 60 POAAGKRNQYOGSAHVSVDKDDLYBFBFDIYNAK--IELPIRKXHLIDLVPRKKG 117
 DB 16 PAAGG---QGGGGIDALTL--DDIIRLLREVTARPGKQVSSBSIRQLCTVSR---- 65
 QY 118 NRLHPKVALLIREAAKSLKQLPNISPVSTAVSQVTVGCGDLHGKLDLLVLRKNGLS 177
 DB 66 -----ALFLSQ-----PMLLELPKIR-----ICDHHQYSDLRLRPEYGGFP 105
 QY 178 SSNPVYFNGDFVDRGKRGLEVLILLSLYLAFPNNAFLNRKMHEDSVNNARYGFIREVS 237
 DB 106 EAM-VIFLGYVDVRGKGSLETTICLLAYKIKYENFPLKRGNECASIRIYGFYBCKR 164
 QY 238 KYPRNKRILAFIDERYRWLPGLSVLSRVLVHGGRS-DSTSLDLIKSIDRGKYSILR 296
 DB 165 RF---NVRIMKVFTEGCFNTLPVALIDDKILCMHGGLSPDLAHLDBRIKNLQR----- 213

QY 297 PVLTDGEPDLKTEWQAIPIIMWSDP-QATMGCVPTLIGAGVWFGPDYTDNFIQRRLSY 355
 DB 214 -----PVDVDPQGLCDLMSDPGKQVQGWGMND-KVSYTTFADKXSEFLQRHDL 265
 QY 356 VIRSHCKPKGHEFPHNDKRTITIFGASNYAIGSNKGYIRLNOLMPHFVQYISAASQT 415
 DB 266 ICRAGQVVDGYEFADRLQVLTITFSAIPNCGEFNDAGAMMSVDETLMSGF-QILMPAEKK 324
 QY 416 KRLSFRQRM 424
 DB 325 QIYGAQKNV 333
 RESULT 35
 ADY22333
 ID ADY22333 standard; protein, 334 AA.
 XX
 AC ADY22333;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Plant full length insert polypeptide seqid 70117.
 XX
 KM plant protectant; plant growth regulant; gene therapy; plant;
 KM recombinant DNA construct; physical array; plant breeding marker;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KM extreme osmotic condition; pathogen tolerance; pest tolerance;
 KM growth rate; cell cycle pathway; disease resistance;
 KM galactomannan production; lignin production; plant growth regulator;
 KM yield; plant growth; plant development; seed oil; protein yield;
 KM protein content.
 XX
 OS Unidentified.
 XX
 PN US2004034888-A1.
 XX
 PD 19-FEB-2004.
 XX
 PF 28-APR-2003; 2003US-00425114.
 XX
 XX 06-MAY-1999; 99US-00304517.
 PR 05-NOV-2001; 2001US-00985678.
 XX
 PA (LIU/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABA/) TABASKA J E.
 PA (CAOV/) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y,
 DR MPI, 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 70117; 15pp; English.
 XX
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous

SQ Sequence 334 AA:

Matches 11

16

66 c

106

165

477

•

AAG30063

17-OCT-20

Protein :
hybridisation

Abstract

1

05-MAR-11

29-MAR-1981

PR 07-JUN-

PR 14-JUN-1967

PR 18-JUN-18-JUN-

PR 18-JUN-18

PR 18-JUN-

22-JUN-
PR

PR 28-JUN-

01-JUL-03

PR 09-JUL-12-JUL-

PR 16-JUL-

PR 19-JUL-

| | | |
|----|-------------|---------------|
| PR | 19-JUL-1999 | 99US-0144333P |
| PR | 19-JUL-1999 | 99US-0144334P |
| PR | 19-JUL-1999 | 99US-0144335P |
| PR | 20-JUL-1999 | 99US-0144332P |
| PR | 20-JUL-1999 | 99US-0144632P |
| PR | 20-JUL-1999 | 99US-0144884P |
| PR | 21-JUL-1999 | 99US-0144814P |
| PR | 21-JUL-1999 | 99US-0145086P |
| PR | 21-JUL-1999 | 99US-0145088P |
| PR | 22-JUL-1999 | 99US-0145085P |
| PR | 22-JUL-1999 | 99US-0145087P |
| PR | 22-JUL-1999 | 99US-0145089P |
| PR | 22-JUL-1999 | 99US-0145152P |
| PR | 23-JUL-1999 | 99US-0145151P |
| PR | 23-JUL-1999 | 99US-0145218P |
| PR | 23-JUL-1999 | 99US-0145224P |
| PR | 26-JUL-1999 | 99US-0145276P |
| PR | 27-JUL-1999 | 99US-0145913P |
| PR | 27-JUL-1999 | 99US-0145918P |
| PR | 28-JUL-1999 | 99US-0145951P |
| PR | 02-AUG-1999 | 99US-0145951P |
| PR | 02-AUG-1999 | 99US-0146388P |
| PR | 03-AUG-1999 | 99US-0146389P |
| PR | 04-AUG-1999 | 99US-0147038P |
| PR | 04-AUG-1999 | 99US-0147204P |
| PR | 05-AUG-1999 | 99US-0147302P |
| PR | 05-AUG-1999 | 99US-0147760P |
| PR | 06-AUG-1999 | 99US-0147703P |
| PR | 06-AUG-1999 | 99US-0147416P |
| PR | 06-AUG-1999 | 99US-0147463P |
| PR | 09-AUG-1999 | 99US-0147935P |
| PR | 10-AUG-1999 | 99US-0148117P |
| PR | 11-AUG-1999 | 99US-0148331P |
| PR | 12-AUG-1999 | 99US-0148331P |
| PR | 13-AUG-1999 | 99US-0148555P |
| PR | 13-AUG-1999 | 99US-0149368P |
| PR | 16-AUG-1999 | 99US-0149368P |
| PR | 17-AUG-1999 | 99US-0149175P |
| PR | 18-AUG-1999 | 99US-0149426P |
| PR | 20-AUG-1999 | 99US-0149722P |
| PR | 20-AUG-1999 | 99US-0149723P |
| PR | 20-AUG-1999 | 99US-0149929P |
| PR | 23-AUG-1999 | 99US-0149930P |
| PR | 23-AUG-1999 | 99US-0149930P |
| PR | 25-AUG-1999 | 99US-0150566P |
| PR | 25-AUG-1999 | 99US-0150884P |
| PR | 26-AUG-1999 | 99US-0151065P |
| PR | 27-AUG-1999 | 99US-0151066P |
| PR | 27-AUG-1999 | 99US-0151080P |
| PR | 30-AUG-1999 | 99US-0151103P |
| PR | 31-AUG-1999 | 99US-0151388P |
| PR | 01-SEP-1999 | 99US-0151530P |
| PR | 01-SEP-1999 | 99US-0152263P |
| PR | 07-SEP-1999 | 99US-0153070P |
| PR | 07-SEP-1999 | 99US-0153758P |
| PR | 13-SEP-1999 | 99US-0154018P |
| PR | 15-SEP-1999 | 99US-0154039P |
| PR | 16-SEP-1999 | 99US-0154779P |
| PR | 20-SEP-1999 | 99US-0155139P |
| PR | 22-SEP-1999 | 99US-0155486P |
| PR | 23-SEP-1999 | 99US-0155569P |
| PR | 24-SEP-1999 | 99US-0156458P |
| PR | 29-SEP-1999 | 99US-0156596P |
| PR | 04-OCT-1999 | 99US-0157117P |
| PR | 05-OCT-1999 | 99US-0157753P |
| PR | 06-OCT-1999 | 99US-0157665P |
| PR | 07-OCT-1999 | 99US-0158032P |
| PR | 08-OCT-1999 | 99US-0158229P |
| PR | 12-OCT-1999 | 99US-0158369P |
| PR | 13-OCT-1999 | 99US-0159293P |
| PR | 13-OCT-1999 | 99US-0159294P |

| | | |
|----|--------------|----------------|
| PR | 13-OCT-1999; | 99US-0159295P. |
| PR | 14-OCT-1999; | 99US-0159329P. |
| PR | 14-OCT-1999; | 99US-0159330P. |
| PR | 14-OCT-1999; | 99US-0159331P. |
| PR | 14-OCT-1999; | 99US-0159637P. |
| PR | 14-OCT-1999; | 99US-0159638P. |
| PR | 18-OCT-1999; | 99US-0159584P. |
| PR | 21-OCT-1999; | 99US-0160741P. |
| PR | 21-OCT-1999; | 99US-0160767P. |
| PR | 21-OCT-1999; | 99US-0160768P. |
| PR | 21-OCT-1999; | 99US-0160770P. |
| PR | 21-OCT-1999; | 99US-0160814P. |
| PR | 21-OCT-1999; | 99US-0160815P. |
| PR | 22-OCT-1999; | 99US-0160980P. |
| PR | 22-OCT-1999; | 99US-0160981P. |
| PR | 22-OCT-1999; | 99US-0160989P. |
| PR | 25-OCT-1999; | 99US-0161404P. |
| PR | 25-OCT-1999; | 99US-0161405P. |
| PR | 25-OCT-1999; | 99US-0161406P. |
| PR | 26-OCT-1999; | 99US-0161359P. |
| PR | 26-OCT-1999; | 99US-0161360P. |
| PR | 26-OCT-1999; | 99US-0161361P. |
| PR | 28-OCT-1999; | 99US-0161920P. |
| PR | 28-OCT-1999; | 99US-0161922P. |
| PR | 28-OCT-1999; | 99US-0161931P. |
| PR | 29-OCT-1999; | 99US-0162142P. |

| | | | | |
|-----------------------|---|--|------------|-------------|
| Query Match | 12.5% | Score 430; | DB 3; | Length 304; |
| Best Local Similarity | 31.2% | Pred. No. 1.9e-31; | | |
| Matches 102; | Conservative 61; | Mismatches 110; | Indels 54; | Gaps 10; |
| QY | 90 | DIYNAKIELPFRKHNDLLIDVFRKKR-----GNRLHPKVALLIREAAKSLKQLPNIS--- | 143 | |
| DB | 8 | DIIRRLVY-----FPRTRPGSGKYH-----LSBG--EIRQLCAVSGKEI | 44 | |
| QY | 144 | -----PSTAVSQQVTCGDLDHGKLDLLVLVLRKGLPSSSNPYVFNQDFVDRKRGLEVL | 199 | |
| DB | 45 | FLQOPNLLLEAPRIKICGDIGHQYSDLLRLPEYGGFPEAN-YLFLGDYVDRGKSLSTI | 103 | |
| QY | 200 | LLLSLYLAFPNAYPLRGRNHDSVMARVCFITEVSKYRNHRKRLIAFIDEVYRWLPL | 259 | |
| DB | 104 | CLLIAYIKIKYKYPENFLLRGNHESASINRIYGFYDECKRRF---NVRLLKIPITDFENCLIPV | 160 | |
| QY | 260 | GSVLNSRVLVYHGFPS-DSTLSLDIKSIDRGKYVSLRPLPLDPSLPDPTKTEMOQIFDIMW | 318 | |
| DB | 161 | AALIDDKLLCMHGSGISPELSLDDIRNAR-----PMDIPSSGLVCDDLLM | 205 | |
| QY | 319 | SDPOATMGCVPTNLRGAGWFGPDVTNPLQRRRLSYVIRSHCKPNQHEFMHDKIITI | 378 | |
| DB | 206 | SDPFGDVGWGMND-RGVSYTFPGADKVAHLEFKHMDILCRAHQVVEDQEFPAERQLVTV | 264 | |
| QY | 379 | FSASNYYAIGSNKQAYIRLNQMLPHF | 405 | |
| DB | 265 | FSAPNYCGRFDMAGAMMSIDSLMCSF | 291 | |
| RESULT 37 | | | | |
| ADT57338 | | | | |
| ID | ADT57338 | standard; protein; 325 AA. | | |
| XX | ADT57338; | | | |
| XX | ADT57338; | | | |
| DB | 13-JAN-2005 | (first entry) | | |
| XX | Plant polypeptide, SEQ ID 7415. | | | |
| XX | Plant; transgenic; cold tolerance; growth rate; drought tolerance; | | | |
| KW | disease resistance; galactomannan production; plant growth regulator; | | | |
| KW | heat tolerance; herbicide tolerance; lignin production; | | | |
| KW | extreme osmotic condition tolerance; lignin production; | | | |
| KW | pest resistance; yield improvement; seed oil yield; seed protein yield. | | | |
| XX | Viridiplantae. | | | |

XX US2004216190-A1.
PN 28-OCT-2004.
XX 18-DEC-2003; 2003US-00739930.
XX 28-APR-2003; 2003US-00424599.
PR 28-APR-2003; 2003US-00425115.
XX (KOVA/) KOVALIC D K.
PA Kovalic DK;
XX MPI; 2004-757369/74.
XX
XX New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.
XX
XX Claim 2; SEQ ID NO 7415; 14pp; English.
XX
XX The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomanan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant protein
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=20040216190.
XX
XX Sequence 325 AA:
SO
Query Match 12.5%; Score 429.5; DB 8; Length 325;
Best Local Similarity 31.6%; Pred. No. 2.4e-11;
Matches 99; Conservative 67; Mismatches 118; Indels 29; Gaps 6;
QY 94 AKLELRKRNHIDLDVFKKRGKGNRLHPYVALLIRBAKSLKOLPNISPVSTANSQOV 153
DB 14 AAVDEVVR-----LVGGRGGRGVQVSEARIRQLCVKRAKVLISQPNL-----LRTHAPV 64
QY 154 TVCGDLHGKDDLLVYHKKGLPSSSNPYVNGDPVDRGKGGLVLLILSLYAPPNV 213
DB 65 KICGDHIGCRVDILRLFDLGGYPTST-VYFLADVDYDKGQSHETICLLAYLTKTPDNT 123

QY 214 PLNRGNHDSVWNAARYGFIREVESKYPNRHKLILAFIDVYKRLPLGSVLSNRVLVHGG 273
DB 124 YILRGNHEDAKINRYGFYDECKRRP--NVRLMKIFCCDFNCLPAAALIDDKILCMHGG 180
QY 274 PS-DSTSLDLKISIDRGKYVSLRPPLTGDGRLDKTEWQOIPDIMSDPOATWGCVPTLL 332
DB 181 LSPELTSLDQIDIER-----PTEIPDYGLCDLMSDSSHOTEGGSESD 225
QY 333 RGAGWFGPDVTNDFLQRRRLSYVIRSHCKENGHEFMNDKIIITIPASNTYAIAGSKNG 392
DB 226 RGVSCTFGADKLVEFLKRNLDLVCRAHQVVDGVEFFAERRLVTIFSA.PNTCGBERDNMG 285
QY 393 AYIRLNNQLMPHF 405
DB 286 ALLSIDESIMCSF 298
RESULT 38
ADTS9253
ID ADTS9253 standard; protein; 326 AA.
XX
XX ADTS9253;
XX
XX 13-JUN-2005 (first entry)
XX
XX Plant polypeptide, SEQ ID 9330.
XX
XX plant; transgenic; cold tolerance; growth rate; drought tolerance;
KW disease resistance; galactomanan production; plant growth regulator;
KW heat tolerance; herbicide tolerance; lignin production;
KW extreme osmotic condition tolerance; pathogens resistance;
KW pest resistance; yield improvement; seed oil yield; seed protein yield.
XX
XX Viridiplantae.
XX
XX US2004216190-A1.
XX
XX 28-OCT-2004.
XX
XX 18-DEC-2003; 2003US-00739930.
XX
XX 28-APR-2003; 2003US-00424599.
XX 28-APR-2003; 2003US-00425115.
XX
XX (KOVA/) KOVALIC D K.
XX
XX Kovalic DK;
XX
XX MPI; 2004-757369/74.
XX
XX New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.
XX
XX Claim 2; SEQ ID NO 9330; 14pp; English.
XX
XX The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomanan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of

QY 382 SNYYAIGSNKGAYIRLNQLMMPH 405
 Db 281 PNYCGEFDNAGALMSIDNSLVCSP 304

RESULT 40
 ADY06935

ID ADY06935 standard; protein; 338 AA.

AC ADY06935;

DT 21-APR-2005 (first entry)

XX Plant full length insert polypeptide seqid 62750.

XX plant protectant; plant growth regulant; gene therapy; plant;
 KM recombinant DNA construct; physical array; plant breeding marker;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KM extreme osmotic condition; pathogen tolerance; pest tolerance;
 KM growth rate; cell cycle pathway; disease resistance;
 KM galactomannan production; lignin production; plant growth regulator;
 KM yield; plant growth; plant development; seed oil; protein yield;
 KM protein content.

XX Unidentified.

XX US2004034888-A1.

PD 19-FEB-2004.

PF 28-APR-2003; 2003US-00425114.

PR 06-MAY-1999; 99US-00304517.

PR 05-NOV-2001; 2001US-00985678.

XX (LIU/) LIU J.

PA (ZHOU/) ZHOU Y.

PA (KOVA/) KOVALIC D K.

PA (SCRE/) SCREEN S E.

PA (TABAS/) TABASKA J E.

PA (CAOY/) CAO Y.

PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

DR WPI; 2004-180133/17.

PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.

PS Claim 1; SEQ ID NO 62750; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.

XX Sequence 338 AA;

Query Match 12.5%; Score 429; DB 8; Length 338;
 Best Local Similarity 31.8%; Pred. No. 2.9e-31;
 Matches 103; Conservative 60; Mismatches 129; Indels 32; Gaps 7;

QY 94 AKIELPIR-----KNIIDLLIDVFRKKGRGRRLHPKVALL---LREAAKLKQLPNI 142
 Db 1 AGISLPSSRSPPGEMDEAAVDDILRLRLRARGRT-PRNAQYDABIRRLCAAAKQVPLS 59
 QY 143 SPVSTAVSQQVTVCGDLHGKLDLVLVHKNGLPSSSNPYVFNGDFVDRGKRLGVLLLL 202
 Db 60 QPVLELEAPRIKICGVHGVQYDGLRLFEVGGTPPPAN-YLRFGLDYVDKSGKSIETICLL 118
 QY 203 LSLVLAFPNAVFLNRGNHEDSVNNARYGFIREVESKYPRNNKRIILAFIDEVYRWLPGLSV 262
 Db 119 LAKIKYPPNPFLLRGNHRCASINRIYGFPEDECKRFP---NVRIMKIFTECFNCLPVAAL 175
 QY 263 LNSRVLIYHGPS-DSTSLDLKSIDRGKYVSLRPPLDGRLDKTEMQQIFDIWMSDP 321
 Db 176 IDDKIFCMHGGUSPDLKSMQIRNIPR-----PVDVPGLLCLDLMSDP 220
 QY 322 QATMGCVPTLKGAGVGFPPDVTDNFLQRRRLSVYIRSHCKPNGHAFMHDKIITIFSA 381
 Db 221 DKEIDRWGENDRGVSYTFGADVAAEFLQKHLDLIGRAHQVVEDGYEFPKQQLVITIFSA 280
 QY 382 SNYYAIGSNKGAYIRLNQLMMPH 405
 Db 281 PNYCGEFDNAGALMSIDNSLVCSP 304

Search completed: January 20, 2006, 19:48:57
 Job time : 139 secs

THIS PAGE BLANK (USPTO)


```
Db 369 IERNR-----QPP--DSGPM-----CDLMSDPQONG-RSISKRGVSCQFPGDYTK 412
Qy 346 NPLQRLRLSVIRSHCKPENGHEFMNDKLTITFSASNTYALGSKGAYIRL-NNQJMPH 404
Db 413 AFLERNLNYIIRSHVKAQYEVAGHGRCVTFSPAPNYCDQMGKASYYHLQGSJLRQ 472
Qy 405 FVOY 408
Db 473 FHOF 476

RESULT 2
US-09-517-779-2
; Sequence 2, Application US/09517779
; Patent No. 6660511
; GENERAL INFORMATION:
; APPLICANT: Luo, Xiang
; APPLICANT: Xu, Xiang
; TITLE OF INVENTION: Cell Cycle Proteins Associated with Rad9, Compositions
; TITLE OF INVENTION: and Methods Of Use
; FILE REFERENCE: A68293/RMS/DAY
; CURRENT APPLICATION NUMBER: US/09/517,779
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-517-779-2
```

```
Query Match 15.8%; Score 543; DB 2; Length 494;
Best Local Similarity 33.4%; Pred. No. 1.1e-43;
Matches 142; Conservative 71; Mismatches 142; Indels 70; Gaps 18;

Qy 11 FLOKTRRRQARRRQRCNMQIFONLEVASBODQALEYKFNDLIKHPQAAGRNOYO 70
Db 90 YIKGYRRAASNNALOK-----FR-----AALRD-----YETVYKVPKHDKAKMKQ 132
Qy 71 -----GSAHV-SVLDDKD---DLVEFPG-DIVAKIELPIKKNHIDL 107
Db 133 ECKNIKOKAFERAIAGDEHKRSVVDLIESWTIDEYSGPLDQKVTIISPMKE----- 188
Qy 108 LIDVFKKKKGNRLHPKCVALLILBAKSLKQLEPNISPVSTANVSQOYTVCGDLHGKLDLL 167
Db 189 LKQWYKDK--KLHRCAYQILVQVKVLSKSTLVEITLKTETKITTVC-DTHGQFYDLL 245
Qy 168 VVLHKNGLPSSSNPY-VFNGDFVDRGKRGLEVLILLSLYLAFPNNAVFLNRGNHEDSVNN 226
Db 246 NIFELANGLSBETNPYFITNGDFVDRGSPVSEVILTLFGKRLYDPHFLIRGNHETDNKN 305
Qy 227 ARYGFIRVESKYPNNHKKILAFIDVYRMLPLGSLNRSVLIHNGF--SDSTSLDLIK 284
Db 306 QLYGFGEVYAKYT--AQWYELFSHFVFWFLAQCINGKLVILMHGGLPSEDOVTLDDIR 362
Qy 285 SIDRGYVAILRPPLDGBPLDKTEWQQLFDIMMSPOATMGCVPTLTLGAGWGPDPVT 344
Db 363 KIERNN-----QPP--DSGPM-----CDLMSDPQONG-RSISKRGVSCQFPGDYT 406
Qy 345 DNFLOHRLSVIRSHCKPENGHEFMNDKLTITFSASNTYALGSKGAYIRL-NNQJMP 403
Db 407 KAFLENNLDYIIRSHVKAQYEVAGHGRCVTFSPAPNYCDQMGKASYYHLQGSJLRQ 466
Qy 404 HFOY 408
Db 467 QHOF 471
```

```
RESULT 3
US-09-270-767-43457
; Sequence 43457, Application US/09270767
; Patent No. 6703491
```

```
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 43457
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43457
```

```
Query Match 15.2%; Score 521.5; DB 2; Length 494;
Best Local Similarity 31.2%; Pred. No. 1.4e-41;
Matches 130; Conservative 72; Mismatches 126; Indels 87; Gaps 14;

Qy 4 NAIIRA-ALFIQKMYRRHQARRM---QRRCNQIFONLEVASBODQALEYKFNDLIK 58
Db 119 SAVKADPAYLKGYYRRAAHMSLGKFKQALCDPEFVAKCR-PNDQDAKIKFTSCNKYIK- 176
Qy 59 MPQAAGRKNOYQGSAAVSLDQKDLVEBFGDIVAKIELPIKKNHIDLIDVFKKGN 118
Db 177 -MRAFERA-----IADVKEKTLSEMYSDMENITIR-----DQY---KGP 212
Qy 119 RLHPRYVALILREAAKSLKQLE-----PNISPVSTAV 149
Db 213 QLEDKGKTL-----KFMKELAKXXXXXXXXXXXXXXXXXXXXXXXXXXXXMBAQPLVDITVPD 266
Qy 150 SQOYTVCGDLHGKLDLVLVHLKNGLPSSSNPYVNGSPVDRGKRGLEVLILLSLYLAF 209
Db 267 EKEFKTICDHIQOQFYDLNNIFELINGLPBKQPYLFPNGDFVDRGSPSVICFTPLFGKLY 326
Qy 210 PNAVLNRGNHEDSVNARVYGFIRVESKYPNNHKKILAFIDVYRMLPLGSLNRSVLI 269
Db 327 PNHFLAAGNHESIMNMQYGTGTGYTAKYT---SAMADIFQVNMPLCHCINQKTLV 383
Qy 270 VHGF--SDSTSLDLIKSIDRGYVSLRPPLDGBPLDKTEWQQLFDIMMSDPQATMG 327
Db 384 MHGGLFSTEDVTLDHIRLERN-----CQPP-----EBGLMCELLMSDPQOMNG- 427
Qy 328 VENTIRGAGWGPVPTDNLQRLSVIRSHCKPENGHEFMNDKLTITFSASNY 384
Db 428 LQSKRGVGIQPGPVTETKFCNNLDYIIRSHVYDMGYEVAHNGKCIYFSASNY 484
```

```
RESULT 4
US-09-248-796A-18327
; Sequence 18327, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18327
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (220), (239)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknc
```



```
/ GENERAL INFORMATION:
/ APPLICANT: Glot, Loic
/ APPLICANT: Mansfield, Traci A.
/ TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
/ FILE REFERENCE: 15966-542
/ CURRENT APPLICATION NUMBER: US/09/538,092
/ CURRENT FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/127,352
/ PRIOR FILING DATE: 1999-04-01
/ PRIOR APPLICATION NUMBER: 60/178,965
/ PRIOR FILING DATE: 2000-02-01
/ NUMBER OF SEQ ID NOS: 1387
/ SOFTWARE: CuraPatSeqFormatter Version 0.9
/ SEQ ID NO 1093
/ LENGTH: 323
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (0)...(0)
/ OTHER INFORMATION: Polypeptide Accession Number P36873
US-09-538-092-1093
```

```
Query Match      11.8%; Score 405.5; DB 2; Length 323;
Best Local Similarity 32.4%; Pred. No. 1.3e-30;
Matches 105; Conservative 58; Mismatches 126; Indels 35; Gaps 11;
```

```
QY 102 KKHIDLLIDVFRKKGRN-----LHPKYV-ALILBAKSLKQLNISPVTAVSQOY 154
DB 6 KNIIDIIORLLEVRSGSKPGKNVQLQBNIRGLCLKSRRIPLSQ-----PILILBAPIK 60
QY 155 VCGDLHGKLDLLVLHKNGLPSSSNPYVFNQDFVDRGKRGLEVLILLISLYAFPNAYF 214
DB 61 ICGDHGOYDILRLFEYGGFPPESSN-YLFLGIDYDRGQSLTICLLAYKIKYENF 119
QY 215 LNHGNEHDSVMNARYGFIREVSKYPRNKHRIIAFIDEYRMLPLGSVLSRVLIVHGPF 274
DB 120 LLEGNHCAISIRIYGFYDECKRRY--NITLMTFTD-CFNCPLAIAYDEKIFCGGL 176
QY 275 S-DSTSLDIKSIDRGKYSILRPLTDSPLDKTEWQOIFDIWMSDP-QATMGCVPTNL 332
DB 177 SPLQSMEOIRRLMR-----PTDVPDQGLCDLWMDPDKDVGWGRNDP 220
QY 333 RGAGVFGPDVTNPLQRRHLSYVIRSHBCKPGHGFMDNKIITIFSASNYAIGSNKG 392
DB 221 RGVSTFFGAEVAKFLKHLDLICRAHQVVEDGYEFPAKQVLTLFSAPIYCGEFDNAG 280
QY 393 AYIRLNNQMLPHFVOYISAASQTK 416
DB 281 AMMSVDETLWCSP-QILKPAKCK 303
```

```
RESULT 8
US-09-167-206-6
/ Sequence 6, Application US/09167206A
/ Patent No. 6476193
/ GENERAL INFORMATION:
/ APPLICANT: Nandabalan, Krishnan
/ APPLICANT: Schulz, Vincent P.
/ APPLICANT: Yang, Melja
/ TITLE OF INVENTION: NIK1 PROTEIN AND NIK1 PROTEIN COMPLEXES
/ FILE REFERENCE: 15966-521 NIK1 protein complexes
/ CURRENT APPLICATION NUMBER: US/09/167,206A
/ CURRENT FILING DATE: 1998-10-06
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 6
/ LENGTH: 330
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-167-206-6
```

```
Query Match      11.7%; Score 403.5; DB 2; Length 330;
```

```
Best Local Similarity 31.7%; Pred. No. 2.1e-30;
Matches 101; Conservative 60; Mismatches 133; Indels 25; Gaps 8;
QY 102 KKHIDLLIDVFRKKGRNRLHPKYVALI---LREAKSLKQLNISPVTAVSQOYVCGD 158
DB 6 KNIIDIIORLLEVRSGSR-PGKNVQLTNEIRGLCLKSRRIPLSQPILILBAPIKICGD 64
QY 159 LHGKLDLLVLHKNGLPSSSNPYVFNQDFVDRGKRGLEVLILLISLYAFPNAYFLNKG 218
DB 65 IHGOYDILRLFEYGGFPPESSN-YLFLGIDYDRGQSLTICLLAYKIKYENFLLNG 123
QY 219 NHEDSVMNARYGFIREVSKYPRNKHRIIAFIDEYRMLPLGSVLSRVLIVHGFS-DS 277
DB 124 NHECASIRIYGFYDECKRRY--NITLMTFTD-CFNCPLAIAYDEKIFCGGLSPDL 180
QY 278 TSLDLIKSIDRGKYSILRPLTDSPLDKTEWQOIFDIWMSDPQATMGCVPTNLRGAGV 337
DB 181 QSMEOIRRLMR-----PTDVPDQGLCDLWMDPDKDVGWGRNDPGRVSR 225
QY 338 WFGPDVTNPLQRRHLSYVIRSHBCKPGHGFMDNKIITIFSASNYAIGSNKQAYIRL 397
DB 226 TFGAEVAKFLKHLDLICRAHQVVEDGYEFPAKQVLTLFSAPIYCGEFDNAGMSV 285
QY 398 NNQMLPHFVOYISAASQTK 416
DB 286 DETLWCSP-QILKPADKCK 303
```

```
RESULT 9
US-09-538-092-873
/ Sequence 873, Application US/09538092
/ Patent No. 6753314
/ GENERAL INFORMATION:
/ APPLICANT: Glot, Loic
/ APPLICANT: Mansfield, Traci A.
/ TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
/ FILE REFERENCE: 15966-542
/ CURRENT APPLICATION NUMBER: US/09/538,092
/ CURRENT FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/127,352
/ PRIOR FILING DATE: 1999-04-01
/ PRIOR APPLICATION NUMBER: 60/178,965
/ PRIOR FILING DATE: 2000-02-01
/ NUMBER OF SEQ ID NOS: 1387
/ SOFTWARE: CuraPatSeqFormatter Version 0.9
/ SEQ ID NO 873
/ LENGTH: 330
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (0)...(0)
/ OTHER INFORMATION: Polypeptide Accession Number P08129
US-09-538-092-873
```

```
Query Match      11.7%; Score 403.5; DB 2; Length 330;
```

```
Best Local Similarity 31.7%; Pred. No. 2.1e-30;
Matches 101; Conservative 60; Mismatches 133; Indels 25; Gaps 8;
```

```
QY 102 KKHIDLLIDVFRKKGRNRLHPKYVALI---LREAKSLKQLNISPVTAVSQOYVCGD 158
DB 6 KNIIDIIORLLEVRSGSR-PGKNVQLTNEIRGLCLKSRRIPLSQPILILBAPIKICGD 64
QY 159 LHGKLDLLVLHKNGLPSSSNPYVFNQDFVDRGKRGLEVLILLISLYAFPNAYFLNKG 218
DB 65 IHGOYDILRLFEYGGFPPESSN-YLFLGIDYDRGQSLTICLLAYKIKYENFLLNG 123
QY 219 NHEDSVMNARYGFIREVSKYPRNKHRIIAFIDEYRMLPLGSVLSRVLIVHGFS-DS 277
DB 124 NHECASIRIYGFYDECKRRY--NITLMTFTD-CFNCPLAIAYDEKIFCGGLSPDL 180
QY 278 TSLDLIKSIDRGKYSILRPLTDSPLDKTEWQOIFDIWMSDPQATMGCVPTNLRGAGV 337
```

Db 181 QSMQIRLRIR-----PTDVPDQGLCDLLMSDPDKDVGQMGENDRGVSP 225
Qy 338 WREGPDYDNDLQHRRLSYVIRSHCKPENGHEFMHNDKIIITPSASYTAIGSKKAYIRL 397
Db 226 TFGAEVVAKFLMHGDDLICRAHQVVDGYEFPAKQOLVTLFSAIPYCGFPAAGAMSV 285
Qy 398 NNQLMHPFVOYISAASQTK 416
Db 286 DETLMCSF-QILKPADKXK 303

RESULT 10
US-09-744-016A-33
; Sequence 33, Application US/09744016A
; Patent No. 6875581
; GENERAL INFORMATION:
; APPLICANT: Dr. Voelkel, Helge
; TITLE OF INVENTION: Method for screening of modulators of calcineurin
; FILE REFERENCE: A34157PCT
; CURRENT APPLICATION NUMBER: US/09/744, 016A
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: EP98113876
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-744-016A-33

Query Match 11.7%; Score 403; DB 2; Length 524;
Best Local Similarity 25.3%; Pred. No. 5.1e-30;
Matches 136; Conservative 78; Mismatches 202; Indels 122; Gaps 16;

Qy 93 NAKIELPIRKNHDLIDVPRKRGKGRNLRHKYVALILREAKSLKQLPNISPVSTAVSQ 152
Db 46 NGKPKVDLKNH-----VKEG-RLBEEVALKTIINGAAILRQ---EKTMIENVAP 92
Qy 153 VTVCGDLHGKLDLIDLVVLRKNGLPSSSNPFYFNGDVFDRGKGLLEVLLLSLYLAPNA 212
Db 93 ITCGDIHQFPLMLFVGVGSPSNTR-YLFLGDYVDRGYSIECVLYIMSLKIHPTK 151
Qy 213 VPLRGNHEDSVMMARVGFIRVYSKYPRNHRKILAFIDEVYKMLPLGVSILNSVLIVHG 272
Db 152 LFLRGNHCRHLTDYFTFKQECRIKY--SEQVYDACMETFDCLPLAALLNQFLCVHG 208
Qy 273 GPG-DSTSLDLSIRGKVSILRPLDGBPLDKTEWQIIPDIMSDPQATMG----- 326
Db 209 GMSPEITSLDDIRKLR-----FTERP-----AFGPVCDLLMSDPSSEDDYGNKXTL 253
Qy 327 --CVPTLRGAGVFGPDYDNDLQHRRLSYVIRSHCKPENGHEFMHNDK-----IIT 378
Db 254 EHTHTHTVAGCSFYGYPAVCEFLQNNLLSTIRAHADAGYRMRKSGOATGFPSELIT 313
Qy 379 FSAASYTAIGSKKAYIRLNNQLMHPFVOYISAASQTKLSFKQRMGIYESSALKEIAVR 438
Db 314 FSAIPYLDVYNNCAAVLKXENNVN----- 337
Qy 439 MRDHRBELDEPRKYPRKDSGYISISHWCKMENVTKGLPRLDLADKLAPGDSQKNY 498
Db 338 -----NIRQFNCSPHY-----WLPNFVDVPTWSLP-----VGEKVEMLV 374
Qy 499 NRTLDLIDTVLLEAADG--MSYMDALYANK-----ASLYAIFNIIDANGSGITLDEPFT 553
Db 375 NVUNICSDDELISDDABEGSTTVRKELIRKIRAIKMARVFSITLQGBSSVUTLGLUP 434
Qy 554 AIDLVAHMPGAVSKAEMLEKCRMDLNGDKVDLNEFLAFLSDLHRKQODENIR 611
Db 435 TGTLPGLVSGKQTE-----TAKQEAABERRAIRGFSIQHKIRISFEAR 480

RESULT 11
US-09-487-558B-274
; Sequence 274, Application US/09487558B
; Patent No. 6949356
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Milne, Todd
; APPLICANT: No. 6949356man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Method for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.130
; CURRENT APPLICATION NUMBER: US/09/487, 558B
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/487, 558
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 446
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 274
; LENGTH: 692
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-487-558B-274

Query Match 11.7%; Score 403; DB 2; Length 692;
Best Local Similarity 28.5%; Pred. No. 8.1e-30;
Matches 120; Conservative 71; Mismatches 166; Indels 64; Gaps 12;

Qy 38 EYASBDQALYKFRPDLIKHNPQAGKRGVYQGSAAHVSLDQKDLVBERGDIYNA--- 94
Db 299 KHIISDDDIENSQSLSNHASMENVDKXN-----NITSKKQPNBEFNDIMOSGN 350
Qy 95 -----KIELPIRKNH-IDLLIDV-FPKRRGNLRHK-----YVALILREAKSLKQLPNT 142
Db 351 KNAKPKFKPIIDETIQKLDAGYAARTRVCLKNNEIIOCTIAKRIPLSQPSLEL 410
Qy 143 SPVSTAVSQVTVCGDLHGKLDLIDLVLRKNGLPSSSNPFYFNGDVFDRGKGLLEVLL 202
Db 411 SP-----PVKIVGVHQYGDLLRLFTKCGFPSSN-YLFLGDYVDRGKQSLERTILL 462
Qy 203 LSLYLAPNAVPLRGNHEDSVMMARVGFIRVYSKYPRNHRKILAFIDEVYKMLPLGSV 262
Db 463 PCYKIKYENPFLLRGNHBCANVTRVGYFDECKKRC--NIKIMKTFID-TPVTLPLAI 519
Qy 263 LNSRVLYIVHGSDSTSLDLSIRGKVSILRPLDGBPLDKTEWQIIPDIMSDPQ 322
Db 520 VAGKIFCHGGLS-----PVLNSMDEIRRV-----RPLDVPDGLINDLMSDPT 565
Qy 322 ATMGCVPTLRGAGVFGPDYDNDLQHRRLSYVIRSHCKPENGHEFMHNDKIIITPSAS 382
Db 566 DDPNEMEDBERGVSCYNNVVALNKPLNKGPFLVCRAHVAVDGYEFPRDRSLVTVFSAP 625
Qy 383 NYTAIGSKKAYIRLNNQLMHPFVOYISAASQTKLSFKQRMGIYESSALKEIAVRMDH 442
Db 626 NYCGBFDWGMGAVMSBGLCSF-----ELDPLDLSALKQVMKKGRQD 669
Qy 443 R 443
Db 670 R 670

RESULT 12
US-09-949-016-7363
; Sequence 7363, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7363
LENGTH: 374
TYPE: PRT
ORGANISM: Human
US-09-949-016-7363

Query Match 11.7%; Score 402.5; DB 2; Length 374;
Best Local Similarity 33.6%; Pred. No. 3.2e-30;
Matches 92; Conservative 50; Mismatches 111; Indels 21; Gaps 6;

QY 144 PVSSTAQSQQVTCGDLHGKLDLVLVLRKNGLPSSNPYVNGDFVDRGKGLGVLTLTL 203
DB 94 PTLLEBAFLKICGDIHQGYDILRLFEYGFPPPSN-YLFLGDYVDRKOSLETICLL 152
QY 204 SLTLAPNAPVFLNRGNHEDSVNNAKYGFLREYSEKTPNNKILAFIDEVYRMPLSGVL 263
DB 153 AYKITYPENFFLLRNGHBCASINRYGFYDECKRKY--NIKLMKFTTD-CFNCLPIAIV 209
QY 264 NSRVLIVHGGS-DESTLDLIRSIDRGKVSILRPLTDSPELDTMWOQIDIMSDPQ 322
DB 210 DEKICCHGSLSPDLOSHQIRIRK-----PTDVPDQGLCDLIMSDPD 254
QY 333 ATMGCVPTLRGAGVWFGPDVTNPLQRRHSYVIRSHCKENGHEFMHDKITITFSAS 382
DB 255 KDVGQMGENDRGVSFTFGAEVAKFLHKKIDLDLICRAHQVVDGYEFPKRLQVLTLFSAF 314
QY 383 NYTAIGSNKAYIRILANQMLPHFVQYISAASQTK 416
DB 315 NYCGEFDNAGAMMSYDETIMCSF-QILKPADRKX 347

RESULT 13
US-09-538-092-1095
Sequence 1095; Application US/09538092
Patent No. 6753314
GENERAL INFORMATION:
APPLICANT: Glot, Loic
APPLICANT: Mannefeld, Traci A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15966-542
CURRENT APPLICATION NUMBER: US/09/538,092
CURRENT FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/127,352
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/178,965
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CuraPatSeqFormatter Version 0.9
SEQ ID NO 1095
LENGTH: 327
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (0)-(0)
OTHER INFORMATION: Polypeptide Accession Number P37140
US-09-538-092-1095

Query Match 11.7%; Score 402; DB 2; Length 327;
Best Local Similarity 31.3%; Pred. No. 2.9e-30;
Matches 96; Conservative 57; Mismatches 128; Indels 26; Gaps 7;
QY 104 HIDLIDVFRKKRGRHLPKYVALILREAAKSL---KQLPNISVSTAVSQQVTCGDL 159
DB 7 NVDSLITLLEVRGR--PGKIVQMTBAVRGLCTIKSREIFLSPILLEBAFLKICGDI 64
QY 160 HGKLDLVLVLRKNGLPSSNPYVNGDFVDRGKGLSVLLLSLYLAFAVAFLNRGN 219
DB 65 HQGYDILRLFEYGFPPPSN-YLFLGDYVDRKOSLETICLLAYKITYPENFFLLNGN 123
QY 220 HEDSVNNAKYGFLREYSEKTPNNKILAFIDEVYRMPLSGVLSRVLYVHGGS-DST 278
DB 124 HECASINNIYGFYDECKRKY--NIKLMKFTTD-CFNCLPIAIVDEKIFCCHGGLSPDLQ 180
QY 279 SLDLIKSIDRGKYVILRPLTDSPELDTMWOQIDIMSDPQATMGCVPNTLRGAGW 338
DB 181 SMEQIRIRIR-----PTDVPDQGLCDLIMSDPDQKDVQMGENDRGVSEFT 225
QY 339 FGPDVTNPLQRRHSYVIRSHCKENGHEFMHDKITITFSASVYTAIGSNKAYIRILN 398
DB 226 FGADVSKFLNRHDLICRAHQVVDGYEFPKRLQVLTLFSAFNYCGSFDNAGAMMSVD 285
QY 399 NQLMPEF 405
DB 286 ETLKCSF 292

RESULT 14
US-09-949-016-6266
Sequence 6266; Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6266
LENGTH: 327
TYPE: PRT
ORGANISM: Human
US-09-949-016-6266

Query Match 11.7%; Score 402; DB 2; Length 327;
Best Local Similarity 31.3%; Pred. No. 2.9e-30;
Matches 96; Conservative 57; Mismatches 128; Indels 26; Gaps 7;
QY 104 HIDLIDVFRKKRGRHLPKYVALILREAAKSL---KQLPNISVSTAVSQQVTCGDL 159
DB 7 NVDSLITLLEVRGR--PGKIVQMTBAVRGLCTIKSREIFLSPILLEBAFLKICGDI 64
QY 160 HGKLDLVLVLRKNGLPSSNPYVNGDFVDRGKGLSVLLLSLYLAFAVAFLNRGN 219
DB 65 HQGYDILRLFEYGFPPPSN-YLFLGDYVDRKOSLETICLLAYKITYPENFFLLNGN 123
QY 220 HEDSVNNAKYGFLREYSEKTPNNKILAFIDEVYRMPLSGVLSRVLYVHGGS-DST 278
DB 124 HECASINNIYGFYDECKRKY--NIKLMKFTTD-CFNCLPIAIVDEKIFCCHGGLSPDLQ 180
QY 279 SLDLIKSIDRGKYVILRPLTDSPELDTMWOQIDIMSDPQATMGCVPNTLRGAGW 338
DB 181 SMEQIRIRIR-----PTDVPDQGLCDLIMSDPDQKDVQMGENDRGVSEFT 225

Db 21 KPRVDILKXKLMK--GRLEBSVAKRIITRGASILRQKKNLDDIDA-----PVTYCGSIHG 74
Qy 162 KLDDLLVTLHKNGLPSSSNPYVFENGDVDRGKGLVLLLLISLYLAEPNAVFLNRGNH 221
Db 75 QPFDLMLKLFEGVGSPLNTR-YLFLGADVDRGFSIECVLYMALKIILPKTLFLLRGNH 133
Qy 222 DSVANNAVGRISREVAEKPRPNHRILAFIDERYNRPLGSLVNAKRVLIYHGGS-DSTSL 260
Db 134 CRHITETFTKQCKCKTY--SERVYDACAQDAFDCILPLAALNQQFLCVHGGLSPKINTL 190
Qy 281 DLKSIDRGKTVSILRPLPLDGEPLDKTEMQOIFDIMSDDPAATWG-----CVPTLIR 333
Db 191 DDIRKIDRFK-----BPP-----AYGMCQILMSDPLEDGGNKKQOEHFTHTYR 235
Qy 334 GAGVWGEQDVTDNFLQHRILSYIYRSHCKPENGHEFMDNK-----IITFSASNTYAI 367
Db 236 GCSYFPYSYPAVCCFLQHNMLILSTRHAEQADAGYMYRKSQYTGPPSLITTFSAFNTLDV 295
Qy 388 GSNKGAYIRLNNQLM-----PH 404
Db 296 YNNKAALVKTENNVMMIRQNCSPH 320

```

RESULT 23
US-09-744-016A-6
Sequence 6, Application US/09744016A
Patent No. 6875581
GENERAL INFORMATION:
APPLICANT: Dr. Voelkel, Helge
TITLE OF INVENTION: Method for screening of modulators of calcineurin
TITLE OF INVENTION: Activity
FILE REFERENCE: A34157PCT
CURRENT APPLICATION NUMBER: US/09/744, 016A
CURRENT FILING DATE: 2001-10-03
PRIOR APPLICATION NUMBER: EP98113876
PRIOR FILING DATE: 1998-07-22
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 510
TYPE: PRT
ORGANISM: Homo sapiens
US-09-744-016A-6

```

[illegible]

```

RESULT 24
US-09-538-092-1283
: Sequence 1283, Application US/09538092
: Patent No. 6753314
: GENERAL INFORMATION:
: APPLICANT: Glot, Lotic
: APPLICANT: Manfield, Traci A.
: TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
: FILE REFERENCE: 15966-542
: CURRENT APPLICATION NUMBER: US/09/538,092
: PRIORITY FILING DATE: 2000-03-29
: PRIOR APPLICATION NUMBER: 60/127,352
: PRIOR FILING DATE: 1999-04-01
: PRIOR APPLICATION NUMBER: 60/178,965
: PRIOR FILING DATE: 2000-02-01
: NUMBER OF SEQ ID NOS: 1387
: SOFTWARE: CurataseqFormatter Version 0.9
: SEQ ID NO 1283
: LENGTH: 521
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURES:
: NAME/KEY: misc feature
: LOCATION: (0)..(0)
: OTHER INFORMATION: Polypeptide Accession Number Q08209
US-09-538-092-1283

```

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 11.4% | Score 391.5; | DB 2; | Length 521; |
| Best Local Similarity | 31.1%; | Pred. NO. 6.6e-29; | | |
| Matches 101; | Conservative 58; | Mismatches 119; | Indels 47; | Gaps 10; |

```

Oy      102 KHNHLLIDVFRKKGNRLHPHYVALILIRBAKSLKOLPNISPVSTAVSQOVTGCHD 165
Db      40 KPRVDILKAHLMKE--GRLEESVALRIITTEGASILIRKCNLLDDA---PVTYCGIHG 93

Oy      162 KLDDLLIVLAKNGLTPESSNRPVYFNGDPYDRGRGLEVLLLSLYLAFPAVNAFLRNGHE 221
Db      94 QPFDLKLKTEEVGQSPANTR-YLPLDGYVDRGTPSIECTLYIMALKIYLPKTLPILRNHE 152

Oy      222 DSVANARGFIREVESAKYPRNHKRIILAFIDEYRNLPLGSVANSRVLIVHGFS-DSTSL 280
Db      153 CRHLETFYFFKQCKCKIY--SERYYDAOMAFDCLPLAALMNGPLCHGSGSPINTL 209

Oy      281 DLIKSIDRKGVYSILRPPLTDSGEPLDKTEWQOIFDIMSDPOATWG-----CVPTLIR 333
Db      210 DDIRKIDRFRK-----EPP-----AYGPMCIDLMSDELDEFGNKTEQHEPHTNVTVR 254

Oy      334 GAGWVGPEPVTNPLFLDRHRLSYVIRSHCKRNGHEPMHNDK-----ITTFASASNYAI 388
Db      255 GCSYFSPYSYAVOCFPLQHNHNLISILRAHRAODAGYRMRYRSQTTGFPSLITTFASAPNYLDV 314

Oy      388 GSNKGAYIRLNNQML-----PH 404
Db      315 YNNKAALVLYKENNVMMIRQFNCSFH 339

```

```

, RESULT 25
, US-09-744-016A-21
, Sequence 21, Application US/09744016A
, Patent No. 687581
, GENERAL INFORMATION:
, APPLICANT: Dr. Voelkel, Helge
, TITLE OF INVENTION: Method for screening of modulators of calcineurin
, FILE REFERENCE: A34157PCT
, CURRENT APPLICATION NUMBER: US/09/744, 016A
, CURRENT FILING DATE: 2001-10-03
, PRIORITY APPLICATION NUMBER: EP98115876
, PRIOR FILING DATE: 1998-07-22
, NUMBER OF SEQ. ID NOS: 35
, SOFTWARE: PatentIn Ver. 2.1
, SEQ ID NO 21

```



```

, CURRENT FILING DATE: 2000-04-14
, PRIOR APPLICATION NUMBER: 60/241,755
, PRIOR FILING DATE: 2000-10-20
, PRIOR APPLICATION NUMBER: 60/237,768
, PRIOR FILING DATE: 2000-10-03
, PRIOR APPLICATION NUMBER: 60/231,498
, PRIOR FILING DATE: 2000-09-08
, NUMBER OF SEQ ID NOS: 207012
, SOFTWARE: FASTSEQ for Windows Version 4.0
, SEQ ID NO: 6461
, LENGTH: 309
, TYPE: PRT
, ORGANISM: Human
US-03-949-016-6461

```

[illegible]

| | Query Match | Similarity | Score | DB 2; | Length | 311; |
|------------|-------------|---|-------|--------|--------|--------|
| Best Local | 92; | Conservative | 63; | Indels | 28; | Gaps 7 |
| Matches | 92; | Conservative | 63; | Indels | 28; | Gaps 7 |
| Qy | 119 | RLHPRYVALILREAAKSLKQLPNIPISVSTAVSQOVTVCGLDHLKQDLDLVTLHKQGLPES | 178 | | | |
| Db | 24 | QINENQVTLCKRAKAILTKESNVQVRC-----PVTVCGLDHQGOFFDLNMLFRIGKSPD | 79 | | | |

```

QY      179  SNPIYFNGNDPDRGKRGLEVLILLISLYIAFNANVFLNKGHTEDSVMANRYGFIKREVESEK 238
Db      80  TN-YLFMDYDYGDSYSEVETLLVALKRYPERITILNGNHSROITOVGYDCLRK 138
QY      239  YPRNKRILIAFIDERYRMPLGASVANSRLVYHGGPSD--TSIDLKSIDRGYVAILRP 297
Db      139  Y--GNANWKYITDLPDYILPFLALVDGQIFCLHGHSPIIDLDHRLALDRLOSVNHEGP 196
QY      298  PLTDSPELDKTEWQOIFDIWMSDP--QATMGCVPTLRLGAGVWFGEDVTDNFLQRIHLSY 355
Db      197  -----MCDLMSDPPDRGGMGSP--RQAGYTFQDDISETFENHANGTL 238
QY      356  VIRSHCEKPNGHGHEMDKILITIFASNYAALGSKGATIRLLNOLMHPFYQY 408
Db      239  VSRAQLVMEAGNWCNCHDRNVYITFSAPNVCYCGNQALAMELDLTLKYGFLOP 291

```

```

QY 106 ULLIDVFPKKKGNRLHPKVALILAEAKSLKQLPNISRVSVAVSQVTCVGDHCKLDD 165
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 40 DLDROIEOLRCEBELLKESEBVKLCAKARILIBESSVQ----RVDSPTVCGDIHQOFTD 95
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 166 LLVLVHLKNGLPSSSNPYVFNQDFVDGRKRGLEVLILLLSLYLAFPAAVFLNRGHDSDVM 225
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 96 LKELFRVQGDVETN-YLFMGDFVDGRGFSVTFILLILNLKRYPRITILIRKHHSRQI 154
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 226 NARYGFIREBSKYPRNHRILAFIDEVYRWLPGLGSVLSRYLIVHGFSDS--TSLDLIK 284
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 155 TVQVGFYDECLARKY--GSVTVWRVCTEIPDYLSLAIIDGKIFCVHGHSFSPISQTLDOIR 212
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 225 SLDKRYVSIARPLTDSRPDLKTEMOQIFDIIMSDPQATMGCVPTTLGAGVTFPPDVT 344
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 213 TIDRKQBPV-----HDG-----PWCDDLMSDPEDTTLWGVSF--RGAGYLLPSQDVV 256
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 345 DNFLGRHLRYVIRSHCEKPNGBHEFMHDKILITIFASVYUVALIGSKGAYIRLNLOLMPH 404
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 257 AOFNAANDIDMICRHQOLMBSGKMHFNETVLITWASAPYCYRCGNVAAILBLDEHLQKD 316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 405 FVOYISAASQTKRLSFKQ 422
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 317 FIFEAAPQETRGISPKK 334
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 31
US-09-949-016-11414

```

: TYPE: PROT
: ORGANISM: Human
US-09-949-016-9733

Query Match      11.1%; Score 383.5; DB 2; Length 312;
Best Local Similarity 33.4%; Pred. No. 1.7e-28;
Matches 97; Conservatve 47; Mismatches 93; Indels 53; Gaps 10

QY      123 KYALLIREAKSLKQLP-----NISVSTAVSQQVTCGDHLGKL 163
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      15 KTV-----EIARLCKYLPENDLAKRLCDVVCDLLBESNVGCVST----PVTVCBGDIHQGF 65
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      164 DDLVLVLAKNKLBPSSNPYVNGFVDRGRKGLVLLLSLLYLAFPAVAVLANKNEDS 223
      : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      66 YDLCELFRRYTGQVDPDIN-YLFMGDFVDRGYSLTFYLLALKKKMPDRITLKGNHRSR 124
      : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      224 VMANRSGFIRESESKYCPRNHRIALAFIDEVYRMPLGSLNRSRLVIVGSPS-DSTSLDL 282
      : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      125 QITQVYGFYDECCQTKY--GNANMARCYCKRYVDMLTYVALIDBQLCVHGSLSPDIKILDQ 182
      : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      283 IKSIDRGKYSILAPPLTDEBPLDKTEWQQLFDIMWSDPQ--ATMGCVPTLRGAGVWFG 340
      : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      183 IRTIERNGEI-----PHKGARCDLVWSDPEVDVFMALSP---RGAGMFLFG 224
      : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      341 PDVTDFNLFQRRRLSYVIRSHCKNGHGFPHNDKIIITTFASNT-YAIGS 389
      : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      225 AKVTNEFVHINNLKLCIGRAQLVHGGYKFMFDEKLVTWWSAPNYCYRCGN 274
      : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 33
US-08-452-722-7
: Sequence 7, Application US/08452722
: Patent No. 5629163
: GENERAL INFORMATION:
: APPLICANT: Scott, John D.
: TITLE OF INVENTION: Calcineurin Inhibitory Compounds and
: TITLE OF INVENTION: Anchoring Protein
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 233 South Wacker Drive, 6300 Sears Tower
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/452,722
: FILING DATE: 30-MAY-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/344,227
: FILING DATE: 23-NOV-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Williams Jr., Joseph A.
: REGISTRATION NUMBER: 38,659
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-474-6300
: TELEFAX: 312-4740448
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 487 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-452-722-7

```


Db 44 VLKNNH-----VKEG-RVDEEIALRLINEGAALIR-----EKTMIIEVADITVCGDI 90
Qy 160 HGLDLDLVLLHKNGLPSSNPVFGNDPVDGRKGLLEVLILLSLYLAFPAVAFPLNRGN 219
Db 91 HGFPLMLKLFVGVGSPANTR-YLFLGDDYVDGYFSIECVLYIMVWKILYPSYLPFLLRGN 149
Qy 220 HEDSVNNAARYGFIREVSKYPRNHKRIIAFIDEVYMLPLGSLVNSRVLIVHGFS-DST 278
Db 150 HECRHILTEYFTFKQCKIKY---SERVYEAACMEAFPSLPLAALLNQOFLCVHGLSPETI 206
Qy 279 SLDLIKSIDRGKVSILRPLDTGEPDLKTEWQIIPDIMSDPQATWG-----CVPT 331
Db 207 TLDDIRLRFRK-----EPP-----AFGPMCDLMSDPSBDFGNEKSGEHSHT 251
Qy 332 LRAGVWFGPDVTDNPLQHRRLSYVIRSHCEKPNHGFPHDNK-----ITIFSASNY 385
Db 252 VRGCSYFYNYPAVCEFLQNNNLISIRAHQAQAGYRMTRKSGTGPSPSLITIFSAPNTL 311
Qy 386 AIGSNKGYIRLNNOLM-----PH 404
Db 312 DVYNNKAVALKYENNVMNIRQFNCSPH 338

RESULT 36

US-08-503-226B-7
; Sequence 7, Application US/08503226B
; Patent No. 5871945
; GENERAL INFORMATION:
; APPLICANT: Lockebie, Robert Owen, et al.
; TITLE OF INVENTION: Calcineurin Inhibitory Compounds and
; TITLE OF INVENTION: Anchoring Protein
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/503,226B
; FILING DATE:
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/404,731
; FILING DATE: 15-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,227
; FILING DATE: 23-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32861
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-4740448
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 487 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-503-226B-7

Query Match 11.0%; Score 380; DB 1; Length 487;
Best Local Similarity 30.9%; Pred. No. 7,7e-28;
Matches 101; Conservative 57; Mismatches 115; Indels 54; Gaps 11;

Qy 100 IRRNHIDLDIVFRKKRGRLHPKYVALIRBAKSLKQLPNISPVSTAVSQOVTGCDI 159
Db 44 VLKNNH-----VKEG-RVDEEIALRLINEGAALIR-----EKTMIIEVADITVCGDI 90
Qy 160 HGLDLDLVLLHKNGLPSSNPVFGNDPVDGRKGLLEVLILLSLYLAFPAVAFPLNRGN 219
Db 91 HGFPLMLKLFVGVGSPANTR-YLFLGDDYVDGYFSIECVLYIMVWKILYPSYLPFLLRGN 149
Qy 220 HEDSVNNAARYGFIREVSKYPRNHKRIIAFIDEVYMLPLGSLVNSRVLIVHGFS-DST 278
Db 150 HECRHILTEYFTFKQCKIKY---SERVYEAACMEAFPSLPLAALLNQOFLCVHGLSPETI 206
Qy 279 SLDLIKSIDRGKVSILRPLDTGEPDLKTEWQIIPDIMSDPQATWG-----CVPT 331
Db 207 TLDDIRLRFRK-----EPP-----AFGPMCDLMSDPSBDFGNEKSGEHSHT 251
Qy 332 LRAGVWFGPDVTDNPLQHRRLSYVIRSHCEKPNHGFPHDNK-----ITIFSASNY 385
Db 252 VRGCSYFYNYPAVCEFLQNNNLISIRAHQAQAGYRMTRKSGTGPSPSLITIFSAPNTL 311
Qy 386 AIGSNKGYIRLNNOLM-----PH 404
Db 312 DVYNNKAVALKYENNVMNIRQFNCSPH 338

RESULT 37

US-08-721-458B-7
; Sequence 7, Application US/08721458B
; Patent No. 6107104
; GENERAL INFORMATION:
; APPLICANT: Lockebie, Robert Owen, et al.
; TITLE OF INVENTION: Calcineurin Inhibitory Compounds
; TITLE OF INVENTION: and Anchoring Protein
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/721,458B
; FILING DATE:
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/404,731
; FILING DATE: 15-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,227
; FILING DATE: 23-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/33276
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-4740448
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 487 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-721-458B-7

PRIOR FILING DATE: 1998-07-22
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 24
LENGTH: 535
TYPE: PRT
ORGANISM: Homo sapiens
US-09-744-016A-24

Query Match 11.0%; Score 380; DB 2; Length 535;
Best Local Similarity 30.9%; Pred. No. 9.1e-28;
Matches 101; Conservative 57; Mismatches 115; Indels 54; Gaps 11;

```
QY 100 IRGNHIDLIDVFRKGRNRLHFKYVALILREAKSLKOLPNISPVSTAVSQQVTCGDL 159
   : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 VLKNHL-----VRKG-RVDEIRIALRIINEGAAILRR---EKTWIEVEAPITVCGDI 111

QY 160 HGLDLDLVVYHKNGLPSSSNPYVFGDPYDRGKRGLEVLILLSLYLAEPNAVPLNRGN 219
   || : || : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 112 HGOFPDLMKLFVGGSPANTR-YLFGDYVDRGYPSIECVLYLWLKILYPSITLFLRGN 170

QY 220 HEDSVNNARYGPIREVESKYPNNHKRIAFIDEVYRWLPFGSVYANGRVLIHGGFS-DST 278
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 HECRHLTBYFTFRQCKIKY---SERVYBACMEAFPSLPLAALLNQFLCVHGLSPRTH 227

QY 279 SLDLKSIDRGKYSILRPPLIDGEPDLDTBWQIIPDIMSDPOATMG-----CVPT 331
   : || : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 228 TLDDIRRLDRFK-----BPP-----AFGPMCDLLMSDPSBDFGNKESQEHFSHT 272

QY 332 LRGAGVWPGPDVTDNFIQRHRLSYVRSHECKPNGHEFMHDK-----ITTFASANTY 385
   : || : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 VRGCSYFYVPVAVCEFLQNNNLISIRAHQAQAGYRMVRSQTTGTFPSLITTFASPNYL 332

QY 386 AIGSNKGAYIRLNNQLM-----PH 404
   : : || : : : : : : : : : : : : : : : : : : : : : : : : :
Db 333 DVTNNKAALVLRKNNVNNIRQFWCSPH 359
```

Search completed: January 20, 2006, 19:53:38
Job time : 48 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2006, 19:52:03 ; Search time 115 Seconds
(without alignments)
2401.610 Million cell updates/sec

Title: US-09-463-733-1

Perfect score: 3442

Sequence: 1 MDENAIRAIFIKWYRRHQ.....VEHDIDPTDCSKVIDPKKS 661

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA Main:
1: /cgn2_6/prodata/1/pubppaa/us07_PUBCOMB.pap:*
2: /cgn2_6/prodata/1/pubppaa/us08_PUBCOMB.pap:*
3: /cgn2_6/prodata/1/pubppaa/us09_PUBCOMB.pap:*
4: /cgn2_6/prodata/1/pubppaa/us10A_PUBCOMB.pap:*
5: /cgn2_6/prodata/1/pubppaa/us10B_PUBCOMB.pap:*
6: /cgn2_6/prodata/1/pubppaa/us11_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|----------------------|---------------------|
| 1 | 3315 | 96.3 | 637 | 6 | US-11-097-143-41259 | Sequence 41259, A |
| 2 | 1305 | 37.9 | 722 | 4 | US-10-369-493-5651 | Sequence 5651, A |
| 3 | 1164.5 | 33.8 | 613 | 4 | US-10-287-226-324 | Sequence 324, App |
| 4 | 1164.5 | 33.8 | 613 | 4 | US-10-287-226-326 | Sequence 326, App |
| 5 | 581 | 16.9 | 483 | 4 | US-10-425-115-347492 | Sequence 347492, A |
| 6 | 581 | 16.9 | 488 | 4 | US-10-425-114-61206 | Sequence 61206, A |
| 7 | 577 | 16.8 | 520 | 6 | US-11-097-143-18312 | Sequence 18312, A |
| 8 | 571 | 16.6 | 500 | 5 | US-10-450-763-40348 | Sequence 40348, A |
| 9 | 570.5 | 16.6 | 483 | 4 | US-10-437-963-195520 | Sequence 195520, A |
| 10 | 570.5 | 16.6 | 485 | 5 | US-10-732-923-9610 | Sequence 9610, App |
| 11 | 570 | 16.6 | 494 | 4 | US-10-043-487-296 | Sequence 296, App |
| 12 | 565 | 16.4 | 556 | 5 | US-10-732-923-9609 | Sequence 9609, App |
| 13 | 562.5 | 16.3 | 482 | 4 | US-10-298-638-27 | Sequence 27, App1 |
| 14 | 562.5 | 16.3 | 485 | 4 | US-10-298-638-10 | Sequence 10, App1 |
| 15 | 561 | 16.3 | 482 | 4 | US-10-424-599-191442 | Sequence 191442, A |
| 16 | 557.5 | 16.2 | 533 | 4 | US-10-298-638-28 | Sequence 28, App1 |
| 17 | 519 | 15.1 | 473 | 4 | US-10-369-493-22595 | Sequence 22595, A |
| 18 | 502.5 | 14.6 | 598 | 4 | US-10-369-493-4091 | Sequence 4091, App1 |
| 19 | 496.5 | 14.4 | 714 | 5 | US-10-450-763-40347 | Sequence 40347, A |
| 20 | 479.5 | 13.9 | 513 | 4 | US-10-369-493-21980 | Sequence 21980, A |
| 21 | 460.5 | 13.4 | 346 | 6 | US-11-097-143-5988 | Sequence 5988, App |
| 22 | 451 | 13.1 | 244 | 4 | US-10-369-493-13007 | Sequence 13007, A |
| 23 | 447.5 | 13.0 | 443 | 4 | US-10-425-115-225312 | Sequence 225312, A |
| 24 | 435.5 | 12.7 | 333 | 4 | US-10-236-699-20 | Sequence 20, App1 |
| 25 | 434.5 | 12.6 | 339 | 4 | US-10-425-114-50385 | Sequence 50385, A |
| 26 | 432.5 | 12.6 | 326 | 4 | US-10-424-599-254277 | Sequence 254277, A |
| 27 | 431.5 | 12.5 | 380 | 4 | US-10-425-114-65099 | Sequence 65099, A |

| | | | | | |
|----|-------|------|-----|------------------------|--------------------|
| 28 | 430.5 | 12.5 | 325 | 5 US-10-739-930-9917 | Sequence 9917, App |
| 29 | 430.5 | 12.5 | 334 | 4 US-10-425-114-64475 | Sequence 64475, A |
| 30 | 430.5 | 12.5 | 334 | 4 US-10-425-114-70117 | Sequence 70117, A |
| 31 | 429.5 | 12.5 | 325 | 5 US-10-739-930-7415 | Sequence 7415, App |
| 32 | 429.5 | 12.5 | 326 | 5 US-10-739-930-9330 | Sequence 9330, App |
| 33 | 429 | 12.5 | 338 | 4 US-10-425-114-62750 | Sequence 62750, A |
| 34 | 429 | 12.4 | 429 | 4 US-10-425-114-66618 | Sequence 66618, A |
| 35 | 428.5 | 12.4 | 428 | 4 US-10-437-963-104697 | Sequence 104697, A |
| 36 | 427.5 | 12.4 | 316 | 4 US-10-369-493-2574 | Sequence 2574, App |
| 37 | 427.5 | 12.4 | 322 | 4 US-10-425-115-29335 | Sequence 29335, A |
| 38 | 427 | 12.4 | 322 | 4 US-10-437-963-111414 | Sequence 111414, A |
| 39 | 427 | 12.4 | 329 | 4 US-10-424-599-194787 | Sequence 194787, A |
| 40 | 426.5 | 12.4 | 322 | 4 US-10-425-115-29336 | Sequence 29336, A |
| 41 | 426.5 | 12.4 | 325 | 4 US-10-437-963-106329 | Sequence 106329, A |
| 42 | 426 | 12.4 | 297 | 4 US-10-425-115-195309 | Sequence 195309, A |
| 43 | 426 | 12.4 | 324 | 4 US-10-425-115-200192 | Sequence 200192, A |
| 44 | 425 | 12.3 | 401 | 4 US-10-369-493-6597 | Sequence 6597, App |
| 45 | 423 | 12.3 | 290 | 4 US-10-425-114-38024 | Sequence 38024, A |

ALIGNMENTS

```
RESULT 1
US-11-097-143-41259
; Sequence 41259, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ASSAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 41259
; LENGTH: 637
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-41259

Query Match      96.3%; Score 3315; DB 6; Length 637;
Best Local Similarity 100.0%; Pred. No. 1.9e+265;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

25 MORRCMOTFONLEYSBODQALTYEPNDLIGMPOAGRKQYOGSAHVSTLDDPDL 84
|||||
1 MORRCMOTFONLEYSBODQALTYEPNDLIGMPOAGRKQYOGSAHVSTLDDPDL 60
|||||
85 VEEFGDIYNAKIELPIRKXHDLLIDVPRKRGGRNRLPPYVALILBEAKSKQLPNISP 144
|||||
61 VEEFGDIYNAKIELPIRKXHDLLIDVPRKRGGRNRLPPYVALILBEAKSKQLPNISP 120
|||||
145 VETAVSQVTVCGDLHGKXDDLVLVHKNGLPSSSNPYFNGDFVDRGRGLEVLILLLS 204
|||||

```

Db      121 VSTAVSQVTCGDLHGKLDLLVLAHKGGLPSSSNPFYFNGDFVDGRGRGLEVLLILLIS 180
Qy      205 LVLAPNAVFLNRGNHDSVMMARVGFIREVESKYPRNKRILAFIDERYRLPLGASVLA 264
Db      181 LVLAPNAVFLNRGNHDSVMMARVGFIREVESKYPRNKRILAFIDERYRLPLGASVLA 240
Qy      265 SHVLVHGSPDSTSLDLIKSIDRGKYVILRPPLTDGEPDLKTEWQOIFDIIMSDPQAT 324
Db      241 SHVLVHGSPDSTSLDLIKSIDRGKYVILRPPLTDGEPDLKTEWQOIFDIIMSDPQAT 300
Qy      325 MCCVPTLRGAGVWFGPDVTNDFLQHRILSVYIRSHCKPNHGFHMDNKITITISASNY 384
Db      301 MCCVPTLRGAGVWFGPDVTNDFLQHRILSVYIRSHCKPNHGFHMDNKITITISASNY 360
Qy      385 VAIGSKGAYIRLNNQMLMHPVOYISAASQTKLS-KQRMGIYESGALKELAVRMRDHD 444
Db      361 VAIGSKGAYIRLNNQMLMHPVOYISAASQTKLS-KQRMGIYESGALKELAVRMRDHD 420
Qy      445 LEDERFRKYDPKDSGYISISHMCKWMENTYKGLPWRLLBDKLA-PGTSQKYNVNTL 504
Db      421 LEDERFRKYDPKDSGYISISHMCKWMENTYKGLPWRLLBDKLA-PGTSQKYNVNTL 480
Qy      505 LDTDVILBAADQMSVMDALYANKASLVAIFNIIDANSGETLDEFEFAIDLVAHMPG 564
Db      481 LDTDVILBAADQMSVMDALYANKASLVAIFNIIDANSGETLDEFEFAIDLVAHMPG 540
Qy      565 AYSKAMLEKCRMDLNGCKYDNLNFELEAFRLSDIHRKQODENIRRSRSTGRPSVAKTA 624
Db      541 AYSKAMLEKCRMDLNGCKYDNLNFELEAFRLSDIHRKQODENIRRSRSTGRPSVAKTA 600
Qy      625 TDPVTLADKISNTLVVHEDIDPTDCSRSKVIDPKKS 661
Db      601 TDPVTLADKISNTLVVHEDIDPTDCSRSKVIDPKKS 637

```

```

RESULT 2
US-10-369-493-5651
/ Sequence 5651, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ PRIOR FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 5651
/ LENGTH: 722
/ TYPE: prt
/ ORGANISM: Caenorhabditis elegans
US-10-369-493-5651

```

```

Query Match      37.9%; Score 1305; DB 4; Length 722;
Best Local Similarity 41.6%; Pred. No. 8 3e-99;
Matches 266; Conservative 117; Mismatches 201; Indels 56; Gaps 13;
Qy      4 NAIRAFIFQWYRRHQAREMQRNCWOIFQNTLEYASQDAE-----LYKFRNDLIK 57
Db      90 STIKSAIILQKWRRCBARLBARRATWQIFALAFAGSDQDLKRWVITLVDFAVDYIR 149
Qy      58 HHPQAAAGRGKQYG--SAIVSVLDD--KQDLYBFRDYNAKIR----- 97
Db      150 AAEENGGKGVNGRNSPLMSALSHYAKSLNDSRGETVYKMLBPTNSPTNVNIDIRNYKGP 209
Qy      98 ---LPIRKVHIDLIDVFRKKGGRNLHPKYVALILBAKSLKQJPNISFVSTAVSQVT 154

```

```

Db      210 TSLSLPDRQVAKMIEAF--KYNKVLHPKCYVLMILHEARKIKRAMPVSRISTISNQVT 267
Qy      155 VCGDLHGKLDLLVLAHKGGLPSSSNPFYFNGDFVDGRGRGLEVLLILLISLVLAPNAV 214
Db      268 ICGDLHGKFDLCLILLYKNGYPSVNPFYFNGDFVDGRGQGSIEVLCVLPALYIVDPMSY 327
Qy      215 LNRGNHDSVMMARVGFIREVESKYPRNKRILAFIDERYRLPLGASVLAHMPG 274
Db      328 LNRGNHDSVMMARVGFIREVESKYPRNKRILAFIDERYRLPLGASVLAHMPG 387
Qy      275 SDSTSLDLIKSIDRGKYVILRPPLTDGEPDLK-----TEWQOIFDIIMSDPQATMGC 327
Db      388 SDSTSLDLIKSIDRGKYVILRPPLTDGEPDLK-----TEWQOIFDIIMSDPQATMGC 447
Qy      328 VNTLRGAGVWFGPDVTNDFLQHRILSVYIRSHCKPNHGFHMDNKITITISASNYVAI 387
Db      448 VNTLRGAGVWFGPDVTNDFLQHRILSVYIRSHCKPNHGFHMDNKITITISASNYVAI 507
Qy      388 GSNKGAAYIR-LNNQMLMHPVOYISAASQTKLS-KQRMGIYESGALKELAVRMRDHD 445
Db      508 GSNKGAAYIR-LNNQMLMHPVOYISAASQTKLS-KQRMGIYESGALKELAVRMRDHD 565
Qy      446 LEDERFRKYDPKDSGYISISHMCKWMENTYKGLPWRLLBDKLA-PGTSQKYNVNTL 504
Db      566 LQKFRINDIEKSGKPLIKMSDQVYRITGLNPLMIALAPKVAITLSEDKVYMYKEDRI 625
Qy      505 LDTDVILBAADQMSVMDALYANKASLVAIFNIIDANSGETLDEFEFAIDLVAHMPG 564
Db      626 AQVGTHAQEKD---IVSLVYHFKSTLTFTLFFRMDKONGGQVSMKEFIDACEVL-----G 677
Qy      565 AYSKAMLEKCRMDLNGCKYDNLNFELEAFRLSDIHRKQODENIRRSRSTGRPSVAKTA 599
Db      678 KYTKRPLQTDYISQIAESIDFNKQGFIDMLBELAFLVLD 717

```

```

RESULT 3
US-10-287-226-324
/ Sequence 324, Application US/10287226
/ Publication No. US20040086875A1
/ GENERAL INFORMATION:
/ APPLICANT: Agee, Michele L.,
/ APPLICANT: Alsbrook, John P.,
/ APPLICANT: Berghs, Constance,
/ APPLICANT: Boldog, Ference,
/ APPLICANT: Burgess, Catherine E.,
/ APPLICANT: Chant, John S.,
/ APPLICANT: Chaudhuri, Amitabha,
/ APPLICANT: Dipippo, Vincent A.,
/ APPLICANT: Edinger, Shlomit R.,
/ APPLICANT: Eisen, Andrew,
/ APPLICANT: Ellerman, Karen,
/ APPLICANT: Gargolli, Basha A.,
/ APPLICANT: Gorman, Linda,
/ APPLICANT: Gerlach, Valerie,
/ APPLICANT: Li, Weizhen,
/ APPLICANT: Kekuda, Ramesh,
/ APPLICANT: Khramtsov, Nikolai,
/ APPLICANT: Li, Li,
/ APPLICANT: Malyskari, Uriel M.,
/ APPLICANT: MacDougall, John R.,
/ APPLICANT: Mezes, Peter S.,
/ APPLICANT: Miller, Charles E.,
/ APPLICANT: Miller, Isabelle,
/ APPLICANT: Ooi, Chean Eng,
/ APPLICANT: Ort, Tatiana,
/ APPLICANT: Padigaru, Muralidhara,
/ APPLICANT: Paturajan, Meera,
/ APPLICANT: Rastelli, Luca,
/ APPLICANT: Rieger, Daniel K.,
/ APPLICANT: Rothenberg, Mark E.,
/ APPLICANT: Shenoy, Suresh G.,
/ APPLICANT: Spaderna, Steven K.,

```

```

; APPLICANT: Spytek, Kimberley A.,
; APPLICANT: Taupier, Jr., Raymond J.,
; APPLICANT: Verniet, Corine A.M.,
; APPLICANT: Zernhusen, Bryan D.,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; PRIOR FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/364,000
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,409
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/364,227
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/334,027
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/331,641
; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 673
; SOFTWARE: Cnaseqblast version 0.1
; SEQ ID NO 324
; LENGTH: 613
; TYPE: PRY
; ORGANISM: Homo sapiens
; US-10-287-226-324

Query Match      33.8%; Score 1164.5; DB 4; Length 613;
Best Local Similarity 39.3%; Pred. No. 2.8e-87;
Matches 247; Conservative 115; Mismatches 200; Indels 67; Gaps 13;

QY      3  ENAIAAIFIOKMYRHHQARREORRCMNIQFONLEYABEQOABLYKFPNDLIKMPQA 62
DB      14  DTSLEBALITQWYRGYKARLAKRQYVATLTTFQSLRYADBOQOMOLSTFSPLEUYTHI 73
QY      63  AGRKNOYOGSAHVSVDLDDK--DLVE--EFGDIVNA-KIELPFRKHIDLLIDVPRKKG 117
DB      74  HHEBELRNGSLESBEDMDRDWDYDSDIVPDSYNGPRLOPFLCTDIDLLLEAFKQO- 132
QY      118  NPLHPKYVALILREAAKSLIKQLPNISPVSTAVSQOYTVCGDLHGKLDLLVYLKNGLPS 177
DB      133  -ILHAYVLEVLFEETKVKLKQMPNFTHIQTSPSKEVTICGDLHGKLDLFLFYKNGLPS 191
QY      178  SNNPVVENDFDYRGKRGLEVLILLISLTLAPNNAVFLRNGHEDSVMMARVCFIREVUS 237
DB      192  ENNPVVFNDPVDKRGNSLEIMILICVSLVYPNDLILRNGHEDPMMNLARGFTTBILH 251
QY      238  KTPRNHRIILAFIDEVYRMLP-----LGSVLSRVLIHVGFSFSDSTSLD 282
DB      252  KYKHGKRILQILBEFYAMLPTETNRDHDGDSKKNKGVTFPA-----HGRIK----- 299
QY      283  IKSIRGKYSILRPLTDEP---LDKTEWQOIFDIWMSDPOATMGCVPTLIRGAGWF 339
DB      300  -----TWGSPTEHILTBHEWQOIIDILWSDRGKNGCFPTCRGGGCVF 342
QY      340  GPDVTDNPLQRRHLSVVRSHCKPBGHPEMDNKITTFESANNYATISNGAYIRLNN 399
DB      343  GPDVTSKILNKYQLKRLRSHCKPBGYBICHDGKVVTFISANNYTBESNGAYIKICS 402
QY      400  QLMPHFVOYISAASQTKRLSF---KORMGIVESSALKELAVMMRDHDELDEPRFYDPR 456
DB      403  GTTPRFQY----QYTKATCFQPLRQRYVTMENSALKILREVIKSKSLTATFAFQLODHR 458
```

```

QY      457  DSGYISISHWCKMENVTKGLPWRLLRDKLAPGTDSDQVNRRTTDLDDTD-VIIRAEA 515
DB      459  KSGKLSVSQWAPCEMIIIDINLPWRSLSSNLVNIIDONGVYVWSSPQNRIRKPVQBAHS 518
QY      516  DGMSVNDALYANKASVVALENIIDANGSEITLDEPETAIDLLVAMPAGVSAEMLEKC 575
DB      519  ---TLVEITRYRSDLBITFNALIDTHSGLSIVSEFRAMKCLFSSHYNVHIDDSQVKNKLA 575
QY      576  RMDLNGDKVDLNEFLFAFRUSDLRKE 604
DB      576  NIMDLNKGSDIPNEFLKAFYV--VARYE 602

RESULT 4
US-10-287-226-326
; Sequence 326, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
; APPLICANT: Agee, Michele L.,
; APPLICANT: Alsobrook, John P.,
; APPLICANT: Berghs, Constance,
; APPLICANT: Boldog, Perence,
; APPLICANT: Burgess, Catherine E.,
; APPLICANT: Chant, John S.,
; APPLICANT: Chaudhuri, Amitabha,
; APPLICANT: DiIppio, Vincent A.,
; APPLICANT: Edinger, Shlomit R.,
; APPLICANT: Elsen, Andrew,
; APPLICANT: Elberman, Karen,
; APPLICANT: Gangolli, Esha A.,
; APPLICANT: German, Linda,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Khrantsev, Nikolai,
; APPLICANT: Li, Li,
; APPLICANT: Maiyankar, Uriel M.,
; APPLICANT: MacDougall, John R.,
; APPLICANT: Mezes, Peter S.,
; APPLICANT: Miller, Charles E.,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Ooi, Chean Eng,
; APPLICANT: Ort, Tatiana,
; APPLICANT: Padigaru, Muralidhara,
; APPLICANT: Patnureajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Rieger, Daniel K.,
; APPLICANT: Rothenberg, Mark E.,
; APPLICANT: Shenoy, Suresh G.,
; APPLICANT: Spaderna, Steven K.,
; APPLICANT: Spytek, Kimberley A.,
; APPLICANT: Taupier, Jr., Raymond J.,
; APPLICANT: Verniet, Corine A.M.,
; APPLICANT: Zernhusen, Bryan D.,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; PRIOR FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/364,000
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,409
```



```

OY 105 IDLLIDVPEKXKKGNRLHAPKCVALLILNEAAKSLKQLGNISPVTAASQOVLVCGDLEKLD 164
Db 181 VAMADEEFKKQK-CIHKYAYQIVLTKTLDLRSMSPSLDVDDVNGCHFTVCGDVIHQYF 238
OY 165 DLLVVLHKKGLPSSSNPYVNGDFVDGRGKGLFVLLLLLSLYLAFPAVAVFLANGNHDSV 224
Db 239 DLIINTFDLNGLSBSEBNPYLPNGDFVDGRGSVSIVILTPAFKCLYFPAKMTIANGNHESKS 298
OY 225 MNARYGFIREVESKYPNRNKRILAFIDBYRWMLPLGSVINSRYLIVHGS-PS-DSTSLDL 282
Db 299 MKRTYGFBESEVNSKLGKFKVELFA---EYFCWPLFPAHVINNKYFVHHGGLFSVDGVGLSD 355
OY 283 IKSIRGKAVSYLIRPPLTGTGBRLDKTEKQOIFEDIMSSDPOATMGCVPNLTIRGAGWFGPD 342
Db 356 IRSIDR-----FCEBP-----EEGIMCELSMDDPQPLQGRGSPK-RGVALLSPGAD 399
OY 343 VTDNFLQRRHSYVIRSHSECKPENGHEFMDNKIITFPSASNTYAIGSNKAYIRLN-NQL 401
Db 400 VYKKFLQENNDLILVRSHEVKEGHEIIEHDGKLIITFSAPNYCDQMGNKGAFTIRTPAEM 459
OY 402 MHPVQY 408
Db 460 KPDIVTF 466

```

```

RESULT 7
US-11-097-143-18312
; Sequence 18312, Application US/11097143
; Publication NO. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CLO00728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18312
; LENGTH: 520
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-18312

Query Match      16.8%; Score 577; DB 6; Length 520;
Best Local Similarity 32.1%; Pred No. 1,1e-38;
Matches 136; Conservative 83; Mismatches 153; Indels 52; Gaps 15

Oy      4  NAIRA-AIPIQKRYRRHQARREM-----QRCNMQIIFONILEVASBODQAEIYKFPNDLIK- 57
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      108 SAVKADPAVILKGYRRRAAHMSIGKFKQALCDFEFAKCR-PNDXDAKLKFTKCNKIYVM 166

Oy      58 -----HMQGAQRKKQYQGSASHVSLDP-KDDVAFBFDIYNAKIEPIRKNHIDL 107
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      167 RAPEBAIANDKREKT-LSSEMYSDMENTITTEDYKGPQLED-----GKVTIKFMKE----- 215

```

```

Qy 108 LIDVPEKKKGNLHPHYVALILREAAKSLKOLPINSPTAVSOQVTCGDHJGKLDLL 167
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 216 LMEHYAOK--RIHRRFAVKIKLCEIDTYWRKOPSJVDITVDEBEKFTICGHIHQFDLM 273
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 168 VLEHKGKGLSSSNPYVFNCDPVDKRGKGLBYLLLSLYLAFPAVAFLANGNHEDSVMA 227
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 274 NIEINGLSBKQPIFLFNGDFVDKRSFSEVCIPLFGFKLLYPHHPFLANGHESIMMQ 333
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 228 RYGFIREVESKYPRNHKRIAFIDEVYMLRPGSVLNSRVLIYHGCF--SDSTSLDLIKS 285
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 334 MYGFTGEVYAKT---SAAADIFTGVFMWMLPLCHCINQKILMHGGLFSEFDVTLDHIRR 390
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 286 IDGKTVSYILRPPLDGBELDKTEWQOIFDIMSDDPQATMGCVNNTLRGAGWFGPVTD 345
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 391 IERN-----COPR-----BERGLMCEIWSPPQOMWG-LQOSKRGVGIQGPDPVTE 434
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 346 NPLQHRHLSYVLRSHCECKNGHREPHNDKILITISASRYVALGSKNKAYIRL--NNQLMPH 404
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 435 KFKQKNNLDYILIRSHBVKDMGYEVAHNKQCTIVPSAPAYCDYTMGNMGAFTITGNNLKPN 494
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 405 FVOY 408
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 495 YKSF 498
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

RESULT 8
US-10-450-763-40348
/ Sequence 40348, Application US/10450763
/ Publication No. US20050196754AI
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc
/ TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
/ FILE REFERENCE: 790CIP3/US
/ CURRENT APPLICATION NUMBER: US/10/450,763
/ PRIOR FILING DATE: 2003-06-11
/ PRIOR APPLICATION NUMBER: PCT/US01/08631
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: 09/540,217
/ PRIOR FILING DATE: 2000-03-31
/ PRIOR APPLICATION NUMBER: 09/649,167
/ PRIOR FILING DATE: 2000-08-23
/ NUMBER OF SEQ ID NOS: 60736
/ SOFTWARE: Custom
/ SEQ ID NO 40348
/ LENGTH: 500
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: DOMAIN
/ LOCATION: (405)..(460)
/ OTHER INFORMATION: Serine/threonine specific protein phosphatases proteins
/ OTHER INFORMATION: domain identified by eMATRIX, accession number BL00125D, p-value=
/ FEATURE:
/ NAME/KEY: DOMAIN
/ LOCATION: (207)..(468)
/ OTHER INFORMATION: Ser/Thr protein phosphatase domain identified by Pfam,
/ OTHER INFORMATION: accession name Sphosphatase, E-value=1.4e-106, Pfam score of 3675
/ NAME/KEY: misc-feature
/ LOCATION: (1)..(500)
/ OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-40348

```

```

Oy      11 FICKWTRRHQAREMOR--RCMKQIFQNTLEYS--BQDAELYKFNFNDLIKHMPOAAGRN 67
      . . . . . : . . . . . : . . . . . : . . . . . : . . . . . :
Ob      99 YINGYVRRASASNNALGKFPFADALRDVEYVKKPHDKAKKMYQDCNNIV--PAAFER-- 154
      . . . . . : . . . . . : . . . . . : . . . . . : . . . . . :

```

QY 68 OYOGSAHV-SVLDKX-----DVEBFG--DIYNAKIEPIRKHIDLIDVERKKRNL 120
DB 155 AIAAGDHKRSVDSLDISEMTIEDYSXKLBEGKTI\$PMKE-----LMQWKXK--KL 208
QY 121 HPRYVALIIREAAKSLKQLEPNISPVSTAVSQOVTGCDLHGKLDLVLVLRKNGLPSSN 180
DB 209 HKKCAQIIVQVKEVLSKSLTVETTLKETEKITVCGDTHGQFYDLNITFELNGLPSSFN 268
QY 181 PVVENDPFVDRGRGLVLLLSLYLAFPNVFLNRGNHDSVMNARGFIREVSKTP 240
DB 269 PTFPNDFVDRGSPSVSVLLTLPFGKLLYPDHFLNRGNHEDVMNQIFGFGEVYAKTY 328
QY 241 RHHKRIAFIDEYRWLPFGSVLNSRVLIVHGF--SDSTSLDLIKSIDRGKTVSLRPP 298
DB 329 ---AGMYELFSVFEWLPFACCGINGVLLMHGGLFSEBDVITLDDIRKIERNR-----QP 380
QY 299 LIDGEPDXTMWOQIFDIMSPOATMGCVPTLRGAGVFGPDVTDFLQRRHLSYIR 358
DB 381 --DSGPM-----CDLMSDPQPONG--RSISKRVTCQFGPDVTYKATLENNLDYIR 429
QY 359 SHECKNGHEFMHNDKIIITIFASANYAIGSNKAYIRL--NQMLPHFOY 408
DB 430 SHEVKAEGTEVAHGRCVTVFSAPNYCDQMGNA\$TYHLQSGDLRPOHOF 480
RESULT 9
US-10-437-963-195520
Sequence 195520, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 195520
LENGTH: 483
TYPE: PRT
ORGANISM: *Oryza sativa*
FEATURES:
OTHER INFORMATION: Clone ID: PAT_MRT4530_91460C.1.pep
US-10-437-963-195520
Query Match 16.6%; Score 570.5; DB 4; Length 483;
Best Local Similarity 34.8%; Pred. No. 3.3e-38;
Matches 145; Conservative 69; Mismatches 154; Indels 49; Gaps 12;
QY 9 AIFIOKQYRRHQBREMQR--RCNMQIFOMLEYASBDOABLYKFYFDLIKMPQAAGRKN 67
DB 77 ARYSKYIRKGAAYLANKGFKKALKFOQVKRISPNDP-----DARRKKECKEKAQV 128
QY 68 OYOGSAHVSLDKDLVBEFGDIVNAKIEL-----PIRKHIDLIDVERKKRNG 117
DB 129 KIRFEBALISVGBEKRKSVASIDYRIIEVBPQYTGPRVNDGTTTLDYFAMLDEFKQK-- 187
QY 118 NRIHPRYVALIIREAAKSLKQLEPNISPVSTAVSQOVTGCDLHGKLDLVLVLRKNGLP 177
DB 188 -CIHKYAYQIVLQVQLRLRSVPSLVDVNPDSHFTVCGDVAGQYFDLLNIFKLGRLS 246
QY 178 SGNPYFNGDFVDRGRGLVLLLSLYLAFPNVFLNRGNHDSVMNARGFIREVSKTP 237
DB 247 EENPYLFNGDFVDRGSPSVSVLLTLPFAFKCLYPTGMYLARGNHESKMKIYGFEBEVS 306
QY 238 KYPRNHRKRIAFID---EYVRWLPGLSVLNSRVLIVHGF--DSTSLDLIKSIDRGKTV 292

DB 307 KUGS-----AFIBAFARFCLPLAHYINNVFVHGGVSDVDKSLDIRAIR----- 356
QY 293 SLRPPDLDGEPDXTMWOQIFDIMSPOATMGCVPTLRGAGVFGPDVTDFLQRRH 352
DB 357 --FREPESG-----LMCEVLMSDPQPOLGRGPRK--RGVGLSFGADVTYKKELOENN 404
QY 353 LSVYRSHCKNGHEFMHNDKIIITIFASANYAIGSNKAYIRL--NQMLPHFOY 408
DB 405 LDVVRSHVDRGEGYIEHHDGLTVFSAPNYCDQMGNA\$ATIRTPALKNISSF 461
RESULT 10
US-10-732-923-9610
Sequence 9610, Application US/10732923
Publication No. US20050108791A1
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
FILE REFERENCE: 38-15(52796)C
CURRENT APPLICATION NUMBER: US/10/732,923
CURRENT FILING DATE: 2003-12-10
PRIOR APPLICATION NUMBER: 10/310,154
PRIOR FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 24149
SEQ ID NO 9610
LENGTH: 485
TYPE: PRT
ORGANISM: *Lycopersicon esculentum*
US-10-732-923-9610
Query Match 16.6%; Score 570.5; DB 5; Length 485;
Best Local Similarity 36.1%; Pred. No. 3.3e-38;
Matches 140; Conservative 55; Mismatches 124; Indels 69; Gaps 12;
QY 24 EMQRRCNQIIFONLEYASBDOABLYKFYFDLIKMPQAAGRKNQYOGSAHVSVLDKXD 83
DB 142 BSGRR---SVADSIDYRS-----VEYBPQYAGARIE----- 169
QY 84 LVEBFGDIVNAKIEPIRKHIDLIDVFRKKGRNLRHPKYVALIIREAASLKQLEPNIS 143
DB 170 ---GDVYTL-----DFYKMLDDPKQKQK--NLHRYAYQIVLQTRRWLRALPSLV 214
QY 144 PVSTAVSQOVTGCDLHGKLDLVLVLRKNGLPSSNPFVFGDPVDRGRGLVLLLT 203
DB 215 DIVPBGKHFTVCGVHGQFPDLNITFELNGLPSEDDNPLFNGDVPDRGSPSEVLLTLP 274
QY 204 SLYLAFPNVFLNRGNHDSVMNARGYFIREVSKYPRNHRKILAFIDEYRWLPFGSVL 263
DB 275 AFKMCQPSAHLARCNHBSKSNKIYGFEBEVSRLSEIFVELFA---EVFCCLPLAHYI 331
QY 264 NSRVLIHGG-PS-DSTSLDLIKSIDRGKTVSILRPPDLDGEPDXTMWOQIFDIMSPO 321
DB 332 NKKVVPVHGGVPSVGVVTKSLDIRAIDR-----FCBPP-----BEGLMCELMSP 376
QY 322 GATMGCVPTLRGAGVFGPDVTDFLQRRHLSYIRSHCKNGHEFMHNDKIIITIFSA 381
DB 377 QPQPRGPRK--RGVGLSFGADVTYKKEFLOENNLDVVRSHVDRGEGYIEHHDGLTVESA 435
QY 382 SNYTAIGSNKAYIRL--NQMLPHFOY 408
DB 436 PNYCDQMGNA\$AFIRFBAVDKMPNITVF 463
RESULT 11
US-10-043-487-296
Sequence 296, Application US/10043487
Publication No. US20030055220A1
GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
APPLICANT: Pierre, LEBRAIN
TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypeptide
TITLE OF INVENTION: mammalian polypeptides

Db 374 QPQPGPGRK-RGVGLSGFQDVTKRFLOENNNLDLVRSHEVXDQGYEIEHDKLITVPSA 432
Qy 382 SNYYAIGSNKGAAYIRLN-NQLMPHFVQY 408
Db 433 PNYCDQMGKGAFFIRFPADEMPNIVTF 460

RESULT 14
US-10-298-638-10
; Sequence 10, Application US/10298638
; Publication No. US20030177527A1
; GENERAL INFORMATION:
; APPLICANT: HARKING, Michel A. et al.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING SIGNAL TRANSDUCTION
; TITLE OF INVENTION: COMPONENTS IN DURABLE AND BROAD-RANGE RESISTANCE
; FILE REFERENCE: 2121-0167P
; CURRENT APPLICATION NUMBER: US/10/298,638
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 10/298,638
; PRIOR FILING DATE: 2002-11-19
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-10-298-638-10

Query Match 16.3%; Score 562.5; DB 4; Length 485;
Best Local Similarity 35.8%; Pred. No. 1.5e-37;
Matches 139; Conservative 55; Mismatches 125; Indels 69; Gaps 12;

Qy 24 EQQRRCNQIFQNLLEYASRQDAELKFFNDLIKHMPOAGRKQNGQSAHVSVDKDD 83
Db 142 EQQR--SVADSIDRS-----VKEBPQYAKIR----- 169
84 LVEERCDIVNAKIELPIRKNIIDLIDVRKKGNLHPKYVALILREAAKSLKOLPNIS 143
Db 170 -----GDVVTL-----DVGKMDLDPKQK--NLHGRVAYQIVLQTRMLRALPSLV 214
Qy 144 PVTSTASQGVYTCGDHLGKLDLVLVHLKNGLPSSNPYFNGDFYDRGKGLVLLIL 203
Db 215 DIVVEGKGFVTCGDVHGQFYDLNLFELNGLPSEBNDPYLFNGDFYDRGSLVLLITLF 274
Qy 204 SLYLAFPAVPLNRGNHDSVNNARVGFIRVESEKPRNKKILAFIDSVYMLPLGSYL 263
Db 275 AKKCMCPSAIHLAGRNHESKSMNKIYGFEGEVRSKLSIFVSLFA---EVPCCLPLAHVI 331
Qy 264 NSRVLIVHG-PS-DSTSLDLIKSIDRGKYSILRPPLTDGSRPLDKTEMQIFDIMS DP 321
Db 332 NKKVPLVHRGLFSVQGVLTSDIRADR-----FCERP-----EGLMCELLMSDP 376
Qy 322 QATMGCVPTLRGAGVFGPVDVTNLFQHRRLSYLIRSHCEKPGNGHEFMHDKIITVPSA 381
Db 377 QPQPGPGRK-RGVGLSGFQDVTKRFLOENNNLDLVRSHEVXDQGYEIEHDKLITVPSA 435
Qy 382 SNYYAIGSNKGAAYIRLN-NQLMPHFVQY 408
Db 436 PNYCDQMGKGAFFIRFPADEMPNIVTF 463

RESULT 15
US-10-424-599-191442
; Sequence 191442, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 191442
LENGTH: 482
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(482)
OTHER INFORMATION: unsure at all Xaa locations
OTHER INFORMATION: Clone ID: PAT_MRT3847_14893C.1.pep
US-10-424-599-191442

Query Match 16.3%; Score 561; DB 4; Length 482;
Best Local Similarity 38.4%; Pred. No. 2e-37;
Matches 135; Conservative 46; Mismatches 119; Indels 52; Gaps 10;

Qy 60 PQAARKNGQGSAAVSVLDDKDLVBERGDIVNAKIELPIRKNIIDLIDVFRKKGNR 119
Db 158 PQYSGAR--IRGDA--VTLDVFKKMIEDF-----KNQKF 187
Qy 120 LHPKTVALLREAAKSLKOLPNISPVSTAVSQGVTCGDHLGKLDLVLVHLKNGLPSSS 179
Db 168 LHKRYAPQIVQTRKXLDALPSLVINHVPDGHFTVCGDVHGQYDLNIFELNGLPSES 247
Qy 180 NPYVNGDFVDRGKGLVLLILSLYLAFPAVPLNRGNHDSVNNARVGFIRVESEK 239
Db 248 NPYLNGDFVDRGSLVLLILSLYLAFKCMSPSAIHLAGRNHESKSMNKIYGFEGEVRSKL 307
Qy 240 PRNKKRIAFIDVRYRMLPLGSVLSRVLVHG-PS-DSTSLDLIKSIDRGKYSILRP 297
Db 308 NETFVELPA---EVPCCPLAHVINEKVFVHGGLSVDGVKVSIDRSINR-----FCER 359
Qy 298 PLTDGSRPLDKTEMQIFDIMS DPQATMGCVPTLRGAGVFGPVDVTNLFQHRRLSYVI 357
Db 360 P-----EGLMCELLMSDPQLPGRGPRK-RGVGLSGAGAVTRFLOENNNLDLVV 408
Qy 368 RSHCEKPGNGHEFMHDKIITVPSASNYAIGSNKGAAYIRLN-NQLMPHFVQY 408
Db 409 RSHCEKPGNGHEFMHDKIITVPSASNYAIGSNKGAAYIRLN-NQLMPHFVQY 460

RESULT 16
US-10-298-638-28
; Sequence 28, Application US/10298638
; Publication No. US20030177527A1
; GENERAL INFORMATION:
; APPLICANT: HARKING, Michel A. et al.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING SIGNAL TRANSDUCTION
; TITLE OF INVENTION: COMPONENTS IN DURABLE AND BROAD-RANGE RESISTANCE
; FILE REFERENCE: 2121-0167P
; CURRENT APPLICATION NUMBER: US/10/298,638
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 10/298,638
; PRIOR FILING DATE: 2002-11-19
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-298-638-28

Query Match 16.2%; Score 557.5; DB 4; Length 533;
Best Local Similarity 39.4%; Pred. No. 4.6e-37;
Matches 127; Conservative 55; Mismatches 109; Indels 31; Gaps 10;

Qy 94 AKIR-LPIRKNIIDLIDVFRKKGNRLHPKTVALLREAAKSLKOLPNISPVSTAVSQ 152

Db 214 ARIEGEVTLDPVYKTMEDPEKOK--TLHRYAYQVLTQRTQLLALPSLVDSVPHGKH 271
Qy 153 VTYGDLHGKLDLVLVHKNGLPSSSNPVVNGDPDRKRGLEVTLLLSLYLAPNA 212
Db 272 ITVCGDVHGQFVYDLNLFELNGLPSEENPYLFNGDFDRSGFSVEILLTLFARKCMPPS 331
Qy 213 VPLRGNHEDSVNARVGFIREVESKYPNNHKLIAFID--EYRMLPLGSLVANSRLV 269
Db 332 IYLARKNHESKSNKXIYGFEGEVRSLKSEK-----FVLLFAVFCYCLPAHYNGKVFY 385
Qy 270 VHGG-ES-DSTSLDLIKSIDRGKYSTLRPPLTDGEBLDTKEWQOIFDIMWSDPQATMGC 327
Db 386 VHGGELFSVDGVCKSLDIRAIDR-----FCEBP-----EGLMCEILLMSDPQPLPGR 430
Qy 328 VPRVTLGAGVWGPEDVTDNLTQHRRLSYVRSHECKNGHEFPHNDKLTITFSASNYAL 387
Db 431 GPRK-RGVGLSFGADVTYKRLQDNMLDLVRSHEVNDGVEYHDGDLTVFSAPYCDQ 489
Qy 388 GSNKGAIVIRLN-NQLMPHFVQY 408
Db 490 MGKGAIFIRFAPDMKPNVTF 511
RESULT 17
US-10-369-493-22595
; Sequence 22595, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052) B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22595
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-22595
Query Match 15.1%; Score 519; DB 4; Length 473;
Best Local Similarity 35.2%; Pred. No. 5.9e-34;
Matches 119; Conservative 66; Mismatches 127; Indels 26; Gaps 8;
Qy 73 AHVSVLDDKDDLVBEFGDIVNAKIELPIRKNHIDLIDVRKKRGNRILHPKYALLIRBA 132
Db 139 ANINIRD--MDIPSDYDYL--LBRQITKBFVEDMKERF--CQGGKPLKRAYSLIRL 191
Qy 133 ASLKLPLNISPVSTIVSQQVTCGLKGLDVLVHKNGLPSSSNPVVNGDFVDRG 192
Db 192 KELLETPSLIDIPVKGDETLVLCGPTHGQYFDLNI FKLHGPPSPNTKYLFGGDFVDVG 251
Qy 193 KRGLEVTLLLSLYLAPNAVPLNRGNHEDSVNARVGFIREVESKYPNNHKLIAFID 252
Db 252 SMTSTVAFTLYAYKLYLPDAVFLNRGNHETDDMKYVFGEGCRSKY--NERTFNIFSE 308
Qy 253 VYRWPLGSLVNSRLVIVHGF--SDSTSLDLIKSIDRGKYSTLRPPLTDGEBLDTKEW 310
Db 309 TSLTLPGLGSLISVYLVHGLPSSDNVTLTDLRNIDRS--KQPGQSG----- 356
Qy 311 QOIFDIMWSDPQATMGCVPNTLRGAGVWGPEDVTDNLTQHRRLSYVRSHECKNGHEFM 370
Db 357 -LMMEMLMTDPOAPRGRSK--RGVGLQGPDPVSKFCEANGIKAVIRSHVADQGYAVE 414
Qy 371 HDNKLTITFSASNYVALGSKNKAIVIRLNQMLMPHFVQY 408

Db 415 HDGCVTFPSAPNYCDSYGNLGAIVIKVEDMELDFHQF 452
RESULT 18
US-10-369-493-4091
; Sequence 4091, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052) B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 4091
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURES:
; NAME/KEY: unsure
; LOCATION: (1)..(598)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-4091
Query Match 14.6%; Score 502.5; DB 4; Length 598;
Best Local Similarity 29.2%; Pred. No. 2e-32;
Matches 141; Conservative 55; Mismatches 148; Indels 139; Gaps 12;
Qy 12 IQKWRBHQARRRMQRQWQIFQNLVYSBDDQAEVYFDFDLIKMPQAGRKQYOG 71
Db 158 VAYIRRRATA-----YAILNPKSAVDFPTCVKIAP----- 189
Qy 72 SAHVSVLDDKD--DLVE-----EFGDIVNAKIELPIRKNHIDLID-- 110
Db 190 -----DNKDAKGLVBECEKIVROLAPPALEVDGELSAAGLDVESNAVDASYGVVL 242
Qy 111 -----VPRKKRGNRLHPKYVALILREAAKSLIKOLPNISPVSTAVSQQVTCG 157
Db 243 EGNEMTQEFIDDMIRFRKRGKLIHKKYVQIILAVANIYVNEPTWVEVDIPEDVQLTVCG 302
Qy 158 DLHG-----KLDLVLVHKNGLPSSSNPVVNGDFPDRKRGLE 197
Db 303 DTHGMPALAMPSPSKWLLTNCEKXQYFDLMELFRLNGFPSPDKRYLLFNGDFVDRGSWSTE 362
Qy 198 VLLLSLYLAPNAVPLNRGNHEDSVNARVGFIREVESK----- 239
Db 363 IALLLYAYKMLRPNNGFIRKGNHETDDMKRYVGFEGEGCHKIKNEKASTLCAHYATSKI 422
Qy 240 -----PRNHKRIAFIDVYRWPLPGSLVANSRLVIVHGF--SDSTSLDLIKS 285
Db 423 PHTPADTSHLLPRTYK-----LFSSEFSLPLATILIGKFFLVHGLFSDDNVTLDDIRK 478
Qy 286 IIRGKYVSLRPPPLTDGEBLDTKEWQOIFDIMWSDPQATMGCVPNTLRGAGVWGPEDVTD 345
Db 479 IDRNHQ--KQPGQSG-----LMMEMLMTDPOAPRGRGRSK-RGVGMQFGPDVTK 524
Qy 346 NTLQHRRLSYVRSHECKNGHEFPHNDKLTITFSASNYVALGSKNKAIVIRLNQMLMPHF 405
Db 525 RCDCKNGLEAIRSHVADQGYAVEYBHDGKCTVFSAPKCDMTEKGAIVINIGPYKLKF 584
Qy 406 VOY 408
Db 585 SQF 587
RESULT 19

US-10-450-763-40347
; Sequence 40347, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/549,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40347
; LENGTH: 714
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (356)..(384)
; OTHER INFORMATION: SERINE/THREONINE PHOSPHATASE FAMILY SIGNATURE domain
; OTHER INFORMATION: identified by eMATRIX, accession number PR00114A, p-value=4.600e-
; OTHER INFORMATION: 25, raw score of 14.44
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (321)..(696)
; OTHER INFORMATION: Ser/Thr protein phosphatase domain identified by Pfam,
; OTHER INFORMATION: accession name Stposphatase, E-value=4.6e-92, Pfam score of 319,
US-10-450-763-40347

Query Match 14.4%; Score 496.5; DB 5; Length 714;
Best Local Similarity 29.2%; Pred. No. 8e-32;
Matches 140; Conservative 62; Mismatches 147; Indels 131; Gaps 15;
QY 41 SBDQAEIKYKFFNDLIKHPQAGRNQYQ-----GSAHV-SYLDXQ- 82
DB 231 SEHEATLTLV---VKKVPHDKDAKMKQKCKYKQAFBRALAGDEHKRSVDSLDI 287
QY 83 ---DIVEEFG--DIVAKIELPIRKNHIDLLIDVFRKKRKNLHPYVALIDREAKSIK 137
DB 288 ESMTEDEYSGPLDEGKVTISPMKE---LMQWYDQK-KLHRKCAQILLVYKEXVL 341
QY 138 QLPNISPVSTAVSQVTVCGDLHGKLDLLVVLHKNGLPSSSNPY----- 183
DB 342 KLSIVETLTKETKLTTCGDTHTGQFYDLNIFELNGLPSEFNTPYVPSFSAEPPSPHLHPQ 401
QY 184 -----FNGDFVDRKKGLEVLTLLSLYLAF 209
DB 402 PPOVCTWEERPLPGKDVTLERRGLAMLHGQGIYFGDVPDRGSRSEVILTLFGKLIY 461
QY 210 PNAVPLNRGNHEDSVNNAARYGFLREYSEKYPNRNKKIILAFIDVYRWMLPLGSLNSRVLI 269
DB 462 PDHFLHLRGNHETDNNGQIYGFGEVYKAYT---AQMYELPSFVEFWMLACINGKXLI 518
QY 270 VNGGF--SDTSLDLKSIDRGKYVSIAPPLTDGSPPLDKTEMQQIDIMSDPQ----- 322
DB 519 MGGGLPSEBDGVTLDTRKIERNR---QPP--DSALPRRMGFVPLARNSGSEPPPHLR 571
QY 323 -----ATMGCVNTL-----RGAGVAFGSDVTDNFIQ 349
DB 572 LBPVHLHGAVMASVTVTWVGEGVTAHTCPYLFNGSISIKRGVSCFGSDVTKAFLE 631
QY 350 RRLSTVIRSHCKPQHEFMHDKIITIFASANYAIGSNKGAAYRL-NNOLMPEFYQY 408
DB 632 ENNLDTIIRSHKAGAGVAVAGRGCVTYFSAFNYCDQGNKASAYIHLQGSDLRPFQHFQ 691
RESULT 20
US-10-369-493-21980

; Sequence 21980, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21980
; LENGTH: 513
; TYPE: PRF
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-21980
Query Match 13.9%; Score 479.5; DB 4; Length 513;
Best Local Similarity 32.3%; Pred. No. 1.2e-30;
Matches 130; Conservative 68; Mismatches 154; Indels 51; Gaps 13;
QY 2 DENAIRAIFIOKWRHRRHQAAREM---QRRCNQIIFQNLVYASBDQAEIKYKFFNDLIK 58
DB 113 DPATKALTLCDRFLRERFRKATGAEAEAKISLCOTLNTSPDANADLANV----- 165
QY 59 MPOAAGRNQYQGSANHVSLDDKDLVSEFGDIVAKIELPIRKNHIDLLIDVFRKKRN 118
DB 166 ----RGPLEFE-----QLYDDKN-----AFKAKIKMNSQERISGVNDLFLK-GK 207
QY 119 RLHPRTVALILREAKSIKQLPNISPV--STAVSQVTVCGDLHGKLDLLVTLHKNGLP 176
DB 208 YLPKTYVAIISHADTLFRQEPSWVLENNSTPVDKISVCDPTHQGFYDVNLTFKFKGV 267
QY 177 GSNNTYVNGPVPDRGKKGLEVLTLLSLYLAFPAVAVLNGNHEDSVNNAARYGIRRYE 236
DB 268 GKPKTYLFGDPVDGNSCEVALLFYCLKILHPNPLFNGNHSDNNKIKYGEDECK 327
QY 237 SKYPNRNKKIILAFIDVYRWMLPLGSLNSRVLIYHGF-SD-STSLDLKSIDRGKYVI 294
DB 328 YKY---SGRITMFRQSFESLPLATLINDVIVHNGLPSPDSATLSDFKNIDR-----F 379
QY 295 LRPPLTDGSPPLDKTEMQQIDIMSDPQATMGCVPTLRGAGVWFQGPVTDNFIQRHRS 354
DB 380 AQP-P-RDQ-----AFMEIWMADPOBANGMGPQ-RGLGHAFGPDIITRFFLRNKLK 428
QY 355 VYIRSHCKPQHEFMHDKIITIFASANYAIGSNKGAAYRL 397
DB 429 KLFPSHRLMGVQVPEQKGLMTVPSAPNYCDQSGNLGSLVHV 471
RESULT 21
US-11-097-143-5988
; Sequence 5988, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932

| | | | | |
|---------------------------|-------|--------------------|------------|-------------|
| Query Match | 13.4% | Score 460.5; | DB 6; | Length 346; |
| Best Local Similarity | 33.8% | Pred. No. 2.6e-29; | | |
| Matches 104; Conservative | 63; | Mismatches 120; | Indels 21; | Gaps 6; |

RESULT 22
HS-10-369-493-13007

Query Match 13.18; Score 451; DB 4; Length 244;

| | | | |
|----|-----|-------------------|-----|
| QY | 397 | LANNOL-----MPH | 404 |
| | : | : | : |
| Db | 222 | IGPELKLEYQVEFAVPH | 238 |

| Year | Best Local Similarity | Pred. No. | 4.5e-28 |
|------|-----------------------|-----------|---------|
| 2000 | 33.5% | 33 | 33 |
| 2001 | 33.5% | 33 | 33 |
| 2002 | 33.5% | 33 | 33 |
| 2003 | 33.5% | 33 | 33 |
| 2004 | 33.5% | 33 | 33 |
| 2005 | 33.5% | 33 | 33 |
| 2006 | 33.5% | 33 | 33 |
| 2007 | 33.5% | 33 | 33 |
| 2008 | 33.5% | 33 | 33 |
| 2009 | 33.5% | 33 | 33 |
| 2010 | 33.5% | 33 | 33 |
| 2011 | 33.5% | 33 | 33 |
| 2012 | 33.5% | 33 | 33 |
| 2013 | 33.5% | 33 | 33 |
| 2014 | 33.5% | 33 | 33 |
| 2015 | 33.5% | 33 | 33 |
| 2016 | 33.5% | 33 | 33 |
| 2017 | 33.5% | 33 | 33 |
| 2018 | 33.5% | 33 | 33 |
| 2019 | 33.5% | 33 | 33 |
| 2020 | 33.5% | 33 | 33 |

1.
 2.
 3.
 4.
 5.
 6.
 7.
 8.
 9.
 10.
 11.
 12.
 13.
 14.
 15.
 16.
 17.
 18.
 19.
 20.
 21.
 22.
 23.
 24.
 25.
 26.
 27.
 28.
 29.
 30.
 31.
 32.
 33.
 34.
 35.
 36.
 37.
 38.
 39.
 40.
 41.
 42.
 43.
 44.
 45.
 46.
 47.
 48.
 49.
 50.
 51.
 52.
 53.
 54.
 55.
 56.
 57.
 58.
 59.
 60.
 61.
 62.
 63.
 64.
 65.
 66.
 67.
 68.
 69.
 70.
 71.
 72.
 73.
 74.
 75.
 76.
 77.
 78.
 79.
 80.
 81.
 82.
 83.
 84.
 85.
 86.
 87.
 88.
 89.
 90.
 91.
 92.
 93.
 94.
 95.
 96.
 97.
 98.
 99.
 100.
 101.
 102.
 103.
 104.
 105.
 106.
 107.
 108.
 109.
 110.
 111.
 112.
 113.
 114.
 115.
 116.
 117.
 118.
 119.
 120.
 121.
 122.
 123.
 124.
 125.
 126.
 127.
 128.
 129.
 130.
 131.
 132.
 133.
 134.
 135.
 136.
 137.
 138.
 139.
 140.
 141.
 142.
 143.
 144.
 145.
 146.
 147.
 148.
 149.
 150.
 151.
 152.
 153.
 154.
 155.
 156.
 157.
 158.
 159.
 160.
 161.
 162.
 163.
 164.
 165.
 166.
 167.
 168.
 169.
 170.
 171.
 172.
 173.
 174.
 175.
 176.
 177.
 178.
 179.
 180.
 181.
 182.
 183.
 184.
 185.
 186.
 187.
 188.
 189.
 190.
 191.
 192.
 193.
 194.
 195.
 196.
 197.
 198.
 199.
 200.
 201.
 202.
 203.
 204.
 205.
 206.
 207.
 208.
 209.
 210.
 211.
 212.
 213.
 214.
 215.
 216.
 217.
 218.
 219.
 220.
 221.

```

; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50385
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURES:
; OTHER INFORMATION: Clone ID: LIB3062-002-C7_F11.pep
US-10-425-114-50385

Query Match 12.6%; Score 434.5; DB 4; Length 339;
Beat Local Similarity 30.1%; Pred. No. 3.6e-27;
Matches 106; Conservative 72; Mismatches 135; Indels 39; Gaps 9;

Oy IKKMPQ-AAGKKNQYSSAVSVLDDDDLVVEFGDIYNAKIELPIRKNIHIDLDIVFRK 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 LRHTPRSPGRK---EGGGGMMTR-----APWGPMEGAAYBMYR-----LVGGRG 46

Oy 56 IKKMPQ-AAGKKNQYSSAVSVLDDDDLVVEFGDIYNAKIELPIRKNIHIDLDIVFRK 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 LRHTPRSPGRK---EGGGGMMTR-----APWGPMEGAAYBMYR-----LVGGRG 46

Oy 115 KRGRRLPKYVALILIRAAKSLKQLPNIPSPVSTAVSQYVVCDDLGKLDLIVLHKNG 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 47 GRQVQLSEATRIQLCVGKRVLSQPNV---LRHAAPVXICGHIHQFDLRLFLPLGG 102

Oy 175 LPSSSNYYVFNQGFVDKRGKGLVLLLLSLYLAFNAVELANGNHDSYMANARYGIR 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 103 YPPAST-YVFLGDVVDKQKSLERITICLLAYKIRYPENITFLGNHDAKINRYGYDE 161

Oy 235 VESYTPNNHKRIILAFIDVYRWLPPLSGVLSRVLYVHGFS-DSTSLDLIKSIDRGKTV 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 162 CKRF--NRYLMKIFSDCFNCLPIALIDDKILCMHGSGSPRTSIDQIKDIER----- 213

Oy 294 ILRPLPLDGRPLDXTENQQLFDIMWSDPQATWGCVPRTLGAQVPGPDVTDNFLQRRL 353
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 214 -----PAKIPDGLLCDLLMDSPSPDGSGMGSDRGVCTFPAADLVVEFLRNLD 263

Oy 354 SYVIRSHCKPENGHEFMHDKIITIFGSASYAYIAISNGKAGYIRLNQMLMPHF 405
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 264 DLIRAGVYEDGYEYFFAQRRLVITISAPNYCGSFDNVGALLGIDSLMSGF 315

RESULT 26
US-10-424-599-254277
; Sequence 254277; Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254277
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURES:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71636C.1.pap
US-10-424-599-254277

Query Match 12.6%; Score 432.5; DB 4; Length 326;
Beat Local Similarity 30.3%; Pred. No. 4.9e-27;
Matches 105; Conservative 69; Mismatches 125; Indels 47; Gaps 9;

Oy 76 SVLDDKDLVEFGDIYNAKIELPIRKNIHIDLDIVFRKGRNLRHPKVALILIRAAKS 135
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 SVLD-----DIIRLLVYTRPG-----KQVQLSEATRIQLCVSREI--- 41

Oy 136 LKQLPNISPVSTAVSQYVVCDDLGKLDLIVLHKNGLPSSNPVFNQGFVDKRGK 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 42 FLQPPNLLLEAPIK---ICGDVHGQYSLLRLFEYGGLPPEAN-YLFLGDVVDKQK 96

```


Best Local Similarity 30.6%; Pred. No. 7.5e-27;
Matches 113; Conservative 67; Mismatches 134; Indels 55; Gaps 13;

```

QY 60 POAAGKNOYQGSAAHVSVDLDDKDLVEBEGDIYNAK--IELPRKNHIDLLIDVPRKKG 117
DB 16 PAAGG---GQGGGIDALL---DDIRRLLEVTARPGKQVLSBSIRQLCTVSR---- 65
QY 118 NRLHPKVALILREBAKSLKQLPNISPVSTAASQVTLGCDLHGKLDLLVLHKNGLPS 177
DB 66 -----AIFLSQ-----PNLELEAPIK-----ICGDHIGQVSDLLRLFEYGGFPP 105
QY 178 SSNPVFNQDFVDRGKGLLEVLILLSLYAFPNNAVFLNKGHEDSVNARVGFIREVES 237
DB 106 EAN-YFLGDYVDRGKQLETTICLLAYKIKYPENFLLRGNECASINRIYGFYDECKR 164
QY 238 KYPBNKHRIIAFIDEVYRMLPGSVLNSRLVIHGGFS-DSTSLDLIKSIDRGKYSILR 296
DB 165 RF---NVRLMKVFTBCEPNTLPVALIIDDKILCMHGLSPDLAHLDEIKMLQR----- 213
QY 297 PPLTDEBPLDKTEWQOIFDIMSDP-QATWGCVPNTLRGAGVFGPDVTDNFLQRRRLSY 355
DB 214 -----PTDVPDQGLICDLIMSDPKQDVQGMGMD-RGVSYTFGADKYSRFLQRRHDL 265
QY 356 VIRSHCKENGHEFPHDNKTIITIFASNTYALGSKNGAYIRLNNOLMPHFVQYISAASOT 415
DB 266 ICRAHQVEDGYEFPADRLQVLTIFSAPNYCGEFDNAGAMMSVDETLMCSF-QLTKPAERK 324
QY 416 KRLSFKQRM 424
DB 325 QIYGAKQNV 333

```

RESULT 30

```

US-10-425-114-70117
/ Sequence 70117, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven B
/ APPLICANT: Tabaska, Jack B
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 70117
/ LENGTH: 334
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: UC-ZMFLMO17236G06_F11.pep
US-10-425-114-70117

```

Query Match 12.5%; Score 430.5; DB 4; Length 334;

Best Local Similarity 30.6%; Pred. No. 7.5e-27;

Matches 113; Conservative 67; Mismatches 134; Indels 55; Gaps 13;

```

QY 60 POAAGKNOYQGSAAHVSVDLDDKDLVEBEGDIYNAK--IELPRKNHIDLLIDVPRKKG 117
DB 16 PAAGG---GQGGGIDALL---DDIRRLLEVTARPGKQVLSBSIRQLCTVSR---- 65
QY 118 NRLHPKVALILREBAKSLKQLPNISPVSTAASQVTLGCDLHGKLDLLVLHKNGLPS 177
DB 66 -----AIFLSQ-----PNLELEAPIK-----ICGDHIGQVSDLLRLFEYGGFPP 105
QY 178 SSNPVFNQDFVDRGKGLLEVLILLSLYAFPNNAVFLNKGHEDSVNARVGFIREVES 237
DB 106 EAN-YFLGDYVDRGKQLETTICLLAYKIKYPENFLLRGNECASINRIYGFYDECKR 164

```

```

QY 238 KYPBNKHRIIAFIDEVYRMLPGSVLNSRLVIHGGFS-DSTSLDLIKSIDRGKYSILR 296
DB 165 RF---NVRLMKVFTBCEPNTLPVALIIDDKILCMHGLSPDLAHLDEIKMLQR----- 213
QY 297 PPLTDEBPLDKTEWQOIFDIMSDP-QATWGCVPNTLRGAGVFGPDVTDNFLQRRRLSY 355
DB 214 -----PTDVPDQGLICDLIMSDPKQDVQGMGMD-RGVSYTFGADKYSRFLQRRHDL 265
QY 356 VIRSHCKENGHEFPHDNKTIITIFASNTYALGSKNGAYIRLNNOLMPHFVQYISAASOT 415
DB 266 ICRAHQVEDGYEFPADRLQVLTIFSAPNYCGEFDNAGAMMSVDETLMCSF-QLTKPAERK 324
QY 416 KRLSFKQRM 424
DB 325 QIYGAKQNV 333

```

RESULT 31

```

US-10-739-930-7415
/ Sequence 7415, Application US/10739930
/ Publication No. US20040216190A1
/ GENERAL INFORMATION:
/ APPLICANT: Kovalic, David K.
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
/ TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
/ FILE REFERENCE: 38-21(53377)B
/ CURRENT APPLICATION NUMBER: US/10/739,930
/ CURRENT FILING DATE: 2003-12-18
/ NUMBER OF SEQ ID NOS: 11088
/ SEQ ID NO 7415
/ LENGTH: 325
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: ZEMA-23APR03-C14584_1.P
US-10-739-930-7415

```

Query Match 12.5%; Score 429.5; DB 5; Length 325;

Best Local Similarity 31.6%; Pred. No. 8.7e-27;

Matches 99; Conservative 67; Mismatches 118; Indels 29; Gaps 6;

```

QY 94 AKIEPRKNHIDLLIDVPRKKGRLHPKVALILREBAKSLKQLPNISPVSTAASQV 153
DB 14 AAVDEVVRR-----LYBGRGRGRQVQMSABIRQLCVAEKQVLLSQPNL-----LRTHAPV 64
QY 154 TVCGDLHGKLDLLVLHKNGLPSSSNPVFNQDFVDRGKGLLEVLILLSLYAFPNNAV 213
DB 65 KICGDHIGQFVDLRLFLDYGYPPTST-YIFLGDYVDRGKQLETTICLLAYKIKYKPDNI 123
QY 214 PLNRGNHEDSVNARVGFIREVESKYPBNKHRIIAFIDEVYRMLPGSVLNSRLVIHGG 273
DB 124 YILRGNEHAKINRYVGFDECKRRF---NVLMKIFCCCFNCLPMAALIIDDKILCMHG 180
QY 274 PS-DSTSLDLIKSIDRGKYSILRPELTGEBPLDKTEWQOIFDIMSDPQATMGCVPTL 332
DB 181 LSPBLTSLDQIKDIR-----PTRIDPYGLICDLIMSDPSDHTBEGMSBD 225
QY 333 RGAGVFGPDVTDNFLQRRRLSYVIRSHCKENGHEFPHDNKTIITIFASNTYALGSKNG 392
DB 226 RGVSCFTFGADKYSRFLQRRHDLVCRAHQVEDGYEFPAERRLVITIFSAPNYCGEFDNAG 285
QY 393 AVIRLNNOLMPHF 405
DB 286 ALISIDESIMCSF 298

```

RESULT 32

```

US-10-739-930-9330
/ Sequence 9330, Application US/10739930
/ Publication No. US20040216190A1
/ GENERAL INFORMATION:
/ APPLICANT: Kovalic, David K.
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH

```

94 AKIELPIR-----KNHIDLIDVFRKRGGRNLHPKYVALI-----LREAAKSLKOLPNI 142

OY 362 SNYYAIGSNKGAYIRLNNQLMPHF 405

```

Db      281 PNYCGEFDNAGALMSIDNSIVCSF 304

RESULT 35
US-10-437-963-104697
; Sequence 104697, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Bardazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 104697
; LENGTH: 428
; TYPE: PRP
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_102005C.1.pep
US-10-437-963-104697

Query Match          12.4%; Score 428.5; DB 4; Length 428;
Best Local Similarity 33.2%; Pred. No. 1.6e-26;
Matches 120; Conservative 48; Mismatches 134; Indels 59; Gaps 9;

Oy      84 LVEERQDIYNAKIELRIRKNHIDLIDVPRKKKGNNRLRPYYVALILREAKSLKQLPNIS 143
         |||::||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      32 LTREWAGLTATLTDWCSRSLPADRLPDV-----LPAAVLORLVLAASVTLRPEPLV 83

Oy      144 PVSTAVSQCVTCGDLHGKLDDLVLVLRKNGLPSSNPYYFNCGDFVVRGRKGLSVLLLL 203
         |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      84 RVDPRPDQVVVVGVGHGQHDVIPILRDAGFPSDDEVFVFNQGDYVDRGMGLFTPTLLL 143

Oy      204 SLYLAFRNAVFLNRGNHDSVMNARYGFIRESKY---PRNKRIIAFTDEVYRWLPL 259
         |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      144 AKKVLLPNSVFLTRGNHESKYCTSYVGFEQSEVNFYKQGCPQVYRKFLRCPED----LPL 199

Oy      260 GSVLMSRVLIIVGGFFDSITSLDIKSIDRG-----KYVSIIRPPILDGEPLDK----- 307
         ::::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      200 ATTINACVTTAHGIGIRGTIIMPSSKRTRYKKGKANKTYANSVDPTLMKLSLDLRLARR 259

Oy      308 -----TEWQOI-----PDIMMSDPQATMGCVNTLRGAGVGFVPDTDNFLQRHRSYYIRS 359
         :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      260 TVLDPRPBGSNLLIPGVDVMSDSPLEBMGLSNKXRGIGLMLGSPDTQQFLYTNNIKLIIRS 319

Oy      360 HECKENGHEFMED-----NKIIITFSASNLYTALG-----SNTGAYTRL 397
         :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      320 HE-GPDARDKRHDLGMDNGYTIIDHHVCCGLITLTFAPDYPOFGASEDRYNNCGAYTVL 378

Oy      398 N 398
Db      379 N 379

RESULT 36
US-10-369-493-2574
; Sequence 2574, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
```

```

APPLICANT: Chen, Xianfeng ;
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIORITY APPLICATION NUMBER: US 60/360,039
PRIORITY FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 2574
LENGTH: 316
TYPE: PRT
ORGANISM: Schizosaccharomyces pombe
FEATURES:
NAME/KEY: unsure
LOCATION: (1)..(316)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-2574

Query Match 12.4%; Score 427.5; DB 4; Length 316;
Best Local Similarity 34.0%; Pred. No. 1,2e-26;
Matches 98; Conservative 53; Mismatches 110; Indels 27; Gaps 7;

QY 144 PVSTRASQOVYCCGDLHGLGDLVLVHNGLPSSNPVFNDDPFDGRGGLGVLLLL 203
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 38 PILELEAPLKIIGIHQYDILRLFEYGGFPEAN-YLFADVDYDRGQSLVLCLL 96
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 204 SLVLAFAVAFLNFGNHDSSVNNARYGFIRESKYPRIHNRILAFIDVYRM,PLGSVL 263
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 97 AYKIKYENPFIILKRNHSCASINRIYGFIDECKRR--NIKLMKLTID-CFNCLPIAII 153
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 264 NSRVLIVHGFS-DETSLDLIKSIDRGKYSILRPETDGEPLDKTEWQOIFDIMSDPQ 322
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 154 DEKIFTHGGGLSPDLNSMDQIQIRK-----PTDVPDTGLCDLWSDPD 198
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 323 ATMGVPTLGLAGGVFPGDVTNFIQRLRLSYVIRSHCKPGRGFMHNKIIITTSAS 382
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 199 KDLTGMGNDRGVSFTFGPDVVSRLHKIDMDLVCAHGVBEYGFESKROLVTLFSAP 258
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 383 NYVAIGSKGAYIRLNNQLMPHFVOYISAASQTKRLSFQRMGIVSS 430
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 259 NYCGFNDAGAMMSYDSEILCSF-QILKPAEK-----KQRYIQGSS 299
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 37
US-10-425-115-293335
; Sequence 293335, Application US/10425115
; Publication NO. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 293335
LENGTH: 322
TYPE: PRT
ORGANISM: Zea mays
FEATURES:
NAME/KEY: unsure
LOCATION: (1)..(322)
OTHER INFORMATION: unsure at all Xaa locations
FEATURES:
OTHER INFORMATION: Clone ID: MRT4577_30602C.1 pep
US-10-425-115-293335

Query Match 12.4%; Score 427.5; DB 4; Length 322;
Best Local Similarity 30.4%; Pred. No. 1,3e-26;

```

Matches 112; Conservative 67; Mismatches 135; Indels 55; Gaps 13;

QY 60 POAAGKNOVQGSAAHVSVDKDDVBERGDIYNAK--IELPRKXNHLIDLVPRKRG 117
DB 5 PAAGG---GGGGGIDALL--DDILRLLEVRTARQVOLSSESRQLCTVSR---54
QY 118 NRLHPKVALILREAAKSLKQLPNIISPVSAVSGQVTVCGDLHKLDDLVLVHKNGLPS 177
DB 55 -----AIFLSQ-----PNLELEAPIK-----ICDIIHQVSDDLRLREYGGFP 94
QY 178 SSNPVYFNGDFVDRGKRGLEVLNLLSLYLAFPNAVFLNKGHDSVMNARYGPIREVS 237
DB 95 EAN-YLFLGDIYVDRGKSLFTICLLAYKIKYENPFILKGNHCASINRIYGFYDECKR 153
QY 238 KYPRNKRILAFIDYVTRMLPLGSLVNSRVLIVHGGS-DSTSLDLIKSIDRGKYVSIIR 296
DB 154 RF---NVRLMKVFTECFNTPALALIDDKILCHGGISPLAHIDELKNIQR-----202
QY 297 PRLTDGEPDLKTEMQOIFDIWMSDP-QATWGCVPNTLRGAGVWGPVTDNFIORHLSY 355
DB 203 -----PTDVPOGLICDLMSDPGKDVQSGKMD-RGVSYTFGADKVEFLQRHDL 254
QY 356 VIRSHCKPNGHGFEMDNKIIITFSASNYAIGSNKGAAYIRLNQMLPHFVQYISAASOT 415
DB 255 ICAHQVVEDGYEFPPADRLQVITIFSAPNYGGERDNAGMMSVDETLKCSF-QILKPAERK 313
QY 416 KRLSFKORM 424
DB 314 XKFGPNKM 322

RESULT 38
US-10-437-963-111414
; Sequence 111414, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 111414
; LENGTH: 322
; TYPE: PRF
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_15395C.1.pep
US-10-437-963-111414

Query Match 12.4%; Score 427; DB 4; Length 322;
Best Local Similarity 30.4%; Pred. No. 1.4e-26;
Matches 114; Conservative 63; Mismatches 136; Indels 62; Gaps 12;

QY 59 MPQAKRNOYQGSAAHVSVDKDDVBERGDIYNAKIELPRKXNHLIDLVPRKRG 118
DB 1 MAAPAGAGGGGGGMDAVLLDD-----ILRR-----LLEVRTARPG-36
QY 119 RLHPKVALILREAAKSLKQLPNIIS---PVSTAVSQVTVCGDLHKLDDLVLVH 171
DB 37 ---KOVQ---SESRQLCTVSRSEIFLSQPNLLEAPIKICDIIHQVSDLLRLPE 88
QY 172 KNGLPSSNPYFNGDFVDRGKRGLEVLNLLSLYLAFPNAVFLNKGHDSVMNARYG 231
DB 89 YGFPPEAN-YLFLGDIYVDRGKSLFTICLLAYKIKYENPFILKGNHCASINRIYGF 147

QY 232 IREVESKPRNKRILAFIDYVTRMLPLGSLVNSRVLIVHGGS-DSTSLDLIKSIDRK 290
DB 148 YDECKRF---NVRLMKVFTECFNCPALALIDDKILCHGGISPLAHIDELKNIQR--202
QY 291 YVSIILAPPLTDGEPDLKTEMQOIFDIWMSDP-QATWGCVPNTLRGAGVWGPVTDNFIQ 349
DB 203 -----PTDVPOGLICDLMSDPGKDVQSGKMD-RGVSYTFGADKVEFLR 248
QY 350 RRLSVIRSHCKPNGHGFEMDNKIIITFSASNYAIGSNKGAAYIRLNQMLPHFVQYI 409
DB 249 KHDLDLCAHQVVEDGYEFPPADRLQVITIFSAPNYGGERDNAGMMSVDETLKCSF-QIL 307
QY 410 SAASOTKRLSFKORM 424
DB 308 KPAERKGFMAASNMK 322

RESULT 39
US-10-424-599-194787
; Sequence 194787, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 194787
; LENGTH: 329
; TYPE: PRF
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_17920C.1.pep
US-10-424-599-194787

Query Match 12.4%; Score 427; DB 4; Length 329;
Best Local Similarity 31.1%; Pred. No. 1.4e-26;
Matches 107; Conservative 68; Mismatches 127; Indels 42; Gaps 10;

QY 105 IDLVDFPKKGNRLHPKVALILREAAKSLKQLPNIIS-----PVSTAVSQVTVCG 157
DB 10 INRLLEV-RGRPGKQV-----LSEA--EIKQLCTVSRDIFLRQPNLLEAPIKIG 59
QY 158 DLHGKLDLVLVHKNGLPSSNPYFNGDFVDRGKRGLEVLNLLSLYLAFPNAVFLNR 217
DB 60 DVHGQVSDLLRLFEYGGLEPPRSN-YLFLGDIYVDRGKSLFTICLLAYKIKYENPFILR 118
QY 218 GHHDSVMNARYGPIREVESKPRNKRILAFIDYVTRMLPLGSLVNSRVLIVHGGS-D 276
DB 119 GHHDCASINRIYGFYDECKRF---NVRLMKVFTECFNCPALALIDDKILCHGGISPB 175
QY 277 STSLDLIKSIDRGKYVSIIRPPLTDGEPDLKTEMQOIFDIWMSDPQATWGCVPNTLRGAG 336
DB 176 LHNLMQIKLPR-----PLEVETGLICDLMSDPSSDIRGGENERGVS 220
QY 337 VWFQPDVTNFIQRLSVIRSHCKPNGHGFEMDNKIIITFSASNYAIGSNKGAAYIR 396
DB 221 YTFGADRVTFEFLQKHDLDLCAHQVVEDGYEFPPARQLVITIFSAPNYGGERDNAGAMMT 280
QY 397 LNNQMLPHFVQYISAASOTKRLSFKORMKIVSSALKELAVMR 440
DB 281 VDETLVCSF-QILKVENKRPSK---GFGSTTVKQSTTRAK 319

RESULT 40
US-10-425-115-283336
; Sequence 283336, Application US/10425115

```
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(5322)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 293336
/ LENGTH: 322
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_30603C.1.pep
US-10-425-115-293336
```

```
Query Match      12.4%; Score 426.5; DB 4; Length 322;
Best Local Similarity 30.1%; Pred. No. 1.5e-26;
Matches 111; Conservative 68; Mismatches 135; Indels 55; Gaps 13;
```

```
QY 60 PQAAGRKNQYOGSAHVSVLDDKDDLYEERPGDIYNAK--TELPIRKHHIDLIDVFAKKGK 117
DB 5 PAAAG---QGGGGIDNAL--DDIRRLLEVRTARPGQVLSSEIRQLCTVSR----- 54
QY 118 NRLHPKYVALILREAAKSLKQLPNISPVSTAVSQOVTVCGLHGLKLDLLVVLHKNGLPS 177
DB 55 -----AIFLSQ-----PNLLRLRAPIK---ICGDIHQYSDLLRLFEYGGFP 94
QY 178 SSNPYFENGDFVDRGKRGLEVLILLSTLYAFPNAVFLNRGNHDSVMNARYGFIREVES 237
DB 95 EAM-YIFLGDYVDRGKQSLFTICLLAYKIKIPENFLRGNHCASINRIYGFDECKR 153
QY 238 KYPNRNKRILAFIDEVYRMPLGSLVNSRVLIHGGFS-DSTSLDLIKSIDRGKYVSIIR 296
DB 154 RF---NVRLMKVYTECFNTLPVAAIDDKILCHHGLSPDLAHLDBIKTLQR----- 202
QY 297 PPLTGGEPJLDKTEWQIPIIMSDP-QATMGCVPTTLRGAQVFGPPDVTDNFLQRHRLSY 355
DB 203 -----PTDVPPQGLCDLWSDPGKDAQGWGMND-RGVSYTFGADKVSFFLQKHLDL 254
QY 356 VIRSHCEKPNGHFPMHDKIITIFASNTYATGSNKGAYIRLNQULMHPFQYISAASQT 415
DB 255 ICAHGVVEDGYEFEPADQLVTIFSAPTYGSEFNDAGAMMSVDETLMCSF-QILKPAERK 313
QY 416 KRLSFKQRM 424
DB 314 NKFMGSNRM 322
```

Search completed: January 20, 2006, 20:03:36
Job time : 118 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2006, 19:52:48 ; Search time 32 Seconds
(without alignments)
209,328 Million cell updates/sec

Title: US-09-463-733-1

Perfect score: 3442
Sequence: 1 MDENAIRAIAFIQWYRHHQ.....VEHDIDPTDCKSKVIDPKXS 661

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA New:
1: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep:*
2: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubppa/PCT_NEW_PUB.pep:*
5: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
7: /cgn2_6/prodata/1/pubppa/US11_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--|
| 1 | 405.5 | 11.8 | 319 | 7 | US-11-109-156-38 Sequence 38, App1 |
| 2 | 405.5 | 11.8 | 323 | 6 | US-10-878-556A-184 Sequence 184, App1 |
| 3 | 405.5 | 11.8 | 323 | 7 | US-11-109-156-37 Sequence 37, App1 |
| 4 | 391.5 | 11.4 | 521 | 7 | US-11-109-156-34 Sequence 34, App1 |
| 5 | 379 | 11.0 | 309 | 7 | US-11-109-156-24 Sequence 24, App1 |
| 6 | 379 | 11.0 | 309 | 7 | US-11-109-156-39 Sequence 39, App1 |
| 7 | 121 | 3.5 | 1614 | 6 | US-10-821-234-903 Sequence 903, App1 |
| 8 | 111 | 3.2 | 548 | 6 | US-10-995-561-810 Sequence 810, App1 |
| 9 | 111 | 3.2 | 567 | 6 | US-10-995-561-813 Sequence 813, App1 |
| 10 | 111 | 3.2 | 559 | 6 | US-10-995-561-812 Sequence 812, App1 |
| 11 | 108 | 3.1 | 709 | 7 | US-11-074-176-158 Sequence 158, App1 |
| 12 | 107.5 | 3.1 | 694 | 7 | US-11-202-516-10 Sequence 10, App1 |
| 13 | 107.5 | 3.1 | 752 | 7 | US-11-202-516-8 Sequence 8, App1 |
| 14 | 105 | 3.1 | 5024 | 6 | US-10-793-626-2964 Sequence 2964, App1 |
| 15 | 104.5 | 3.0 | 1070 | 7 | US-11/062 Sequence 4, App1 |
| 16 | 104.5 | 3.0 | 1095 | 7 | US-11/062 Sequence 4, App1 |
| 17 | 103 | 3.0 | 270 | 6 | US-10-118-590-18 Sequence 18, App1 |
| 18 | 103 | 3.0 | 1765 | 6 | US-10-055-877-140 Sequence 140, App1 |
| 19 | 102 | 3.0 | 149 | 6 | US-10-821-234-1535 Sequence 1535, App1 |
| 20 | 102 | 3.0 | 149 | 7 | US-11-124-368A-268 Sequence 268, App1 |
| 21 | 102 | 3.0 | 149 | 7 | US-11-124-368A-269 Sequence 269, App1 |
| 22 | 102 | 3.0 | 171 | 6 | US-10-821-234-935 Sequence 935, App1 |
| 23 | 101.5 | 2.9 | 148 | 7 | US-11-065-943-46 Sequence 46, App1 |
| 24 | 101 | 2.9 | 257 | 6 | US-10-118-590-16 Sequence 16, App1 |
| 25 | 100.5 | 2.9 | 1548 | 7 | US-11-108-172-1095 Sequence 1095, App1 |

| | | | | | |
|----|------|-----|------|---|--|
| 26 | 99 | 2.9 | 276 | 6 | US-10-467-657-2502 Sequence 2502, App1 |
| 27 | 98.5 | 2.9 | 746 | 7 | US-11-074-176-314 Sequence 314, App1 |
| 28 | 98.5 | 2.9 | 749 | 7 | US-11-074-176-54 Sequence 54, App1 |
| 29 | 98 | 2.8 | 1538 | 6 | US-10-995-561-772 Sequence 772, App1 |
| 30 | 98 | 2.8 | 5335 | 6 | US-10-995-561-777 Sequence 777, App1 |
| 31 | 98 | 2.8 | 5406 | 6 | US-10-995-561-774 Sequence 774, App1 |
| 32 | 98 | 2.8 | 5415 | 6 | US-10-995-561-779 Sequence 779, App1 |
| 33 | 98 | 2.8 | 5464 | 6 | US-10-995-561-775 Sequence 775, App1 |
| 34 | 98 | 2.8 | 5935 | 6 | US-10-995-561-776 Sequence 776, App1 |
| 35 | 97.5 | 2.8 | 467 | 7 | US-11-156-084-115 Sequence 115, App1 |
| 36 | 97.5 | 2.8 | 613 | 6 | US-10-131-826A-190 Sequence 190, App1 |
| 37 | 97 | 2.8 | 216 | 6 | US-10-118-590-2 Sequence 2, App1 |
| 38 | 97 | 2.8 | 216 | 6 | US-10-118-590-6 Sequence 6, App1 |
| 39 | 97 | 2.8 | 814 | 6 | US-10-878-556A-161 Sequence 161, App1 |
| 40 | 97 | 2.8 | 1184 | 7 | US-11-115-639-50 Sequence 50, App1 |
| 41 | 97 | 2.8 | 4384 | 6 | US-10-821-234-1120 Sequence 1120, App1 |
| 42 | 96.5 | 2.8 | 1145 | 6 | US-10-793-626-1432 Sequence 1432, App1 |
| 43 | 95.5 | 2.8 | 2897 | 6 | US-10-499-715-2 Sequence 2, App1 |
| 44 | 95 | 2.8 | 225 | 6 | US-10-118-590-30 Sequence 30, App1 |
| 45 | 95 | 2.8 | 252 | 6 | US-10-118-590-22 Sequence 22, App1 |

ALIGNMENTS

RESULT 1
US-11-109-156-38
; Sequence 38, Application US/1109156
; Publication No. US20050250144A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Oca
; APPLICANT: Takao Isogai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Wakamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-ichi Funahashi
; APPLICANT: Chiaki Senoo
; APPLICANT: Jun-ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/11/109,156
; PRIOR FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: US/10/060,065
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-109-156-38

```
Query Match          11.8%; Score 405.5; DB 7; Length 319;
Best Local Similarity 32.4%; Pred. No. 2.7e-24;
Matches 105; Conservative 58; Mismatches 126; Indels 35; Gaps 11;

QY 102 KKHIDLLIVFRKRGNR-----LHPKTV-ALLIREAKSKIKOLPNISPVSTAVSQOVT 154
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 2 KLNIDSIIORLLEVRGSKRGKVVQLOENIRIGLCLKSREIFLSQ-----PILBLEAPLK 56

QY 155 VCGDHLHGKLDLLVVLHKNGLPSSSNPVYFNGDFVDRGKRGLEVVLLLSLYLAPNNAV 214
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 57 ICGDHLHGQYDDLRLREFYGGPPPSN-YLFLGDYVDRGKOSLETICLLAYKIKYKPENF 115

QY 215 LNRGNHDSVMNARVGFIREVESKYPRNHKRIIAFIDERYRWPLGVSYVNSKVLVHGCF 274
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 116 LNRGNHSCASINRIYGFYDECKRRY--NIKLMKTFTD-CFNCPLPIAAYDEKIFCCHGGL 172

QY 275 S-DSTSLDLIKSIDRGKTVSILRPPLTDGEPIDKTEWQOIPDIMSDP-QATMGCVPTLT 332
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 173 SPDLQSMGQIRIRIM-----PTDVPOGLLCDLMSDPDKOVLGNGEND- 216

QY 333 RGAGWFGPDYTDNPLQHRRLSVIRSHCKRNGHSPMDNKIITTFASANTYVAGSNG 392
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 217 RGVSTFGAEVVAKFIAHKDLIDLCRAHQVVEDGHEFAKQDLVTLFSAPTYCGSEPDNAG 276

QY 393 AYIRLNQMLMPHFVQYISAASQTK 416
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 277 AMMSVDETLMCSF-QILKPAEKKK 299

RESULT 2
US-10-878-556A-184
/ Sequence 184, Application US/10878556A
/ Publication No. US2005026399A1
/ GENERAL INFORMATION:
/ APPLICANT: Hoffmann La-Roche Inc.
/ TITLE OF INVENTION: HCV regulated protein expression
/ FILE REFERENCE: 21762
/ CURRENT APPLICATION NUMBER: US/10/878,556A
/ CURRENT FILING DATE: 2004-06-28
/ NUMBER OF SEQ ID NOS: 199
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 184
/ LENGTH: 323
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: sw_hum/pp1g_human
/ DATABASE ENTRY DATE: 1994-06-01
/ US-10-878-556A-184

Query Match          11.8%; Score 405.5; DB 6; Length 323;
Best Local Similarity 32.4%; Pred. No. 2.8e-24;
Matches 105; Conservative 58; Mismatches 126; Indels 35; Gaps 11;

QY 102 KKHIDLLIVFRKRGNR-----LHPKTV-ALLIREAKSKIKOLPNISPVSTAVSQOVT 154
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 6 KLNIDSIIORLLEVRGSKRGKVVQLOENIRIGLCLKSREIFLSQ-----PILBLEAPLK 60

QY 155 VCGDHLHGKLDLLVVLHKNGLPSSSNPVYFNGDFVDRGKRGLEVVLLLSLYLAPNNAV 214
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 61 ICGDHLHGQYDDLRLREFYGGPPPSN-YLFLGDYVDRGKOSLETICLLAYKIKYKPENF 119

QY 215 LNRGNHDSVMNARVGFIREVESKYPRNHKRIIAFIDERYRWPLGVSYVNSKVLVHGCF 274
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 120 LNRGNHSCASINRIYGFYDECKRRY--NIKLMKTFTD-CFNCPLPIAAYDEKIFCCHGGL 176

QY 275 S-DSTSLDLIKSIDRGKTVSILRPPLTDGEPIDKTEWQOIPDIMSDP-QATMGCVPTLT 332
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 177 SPDLQSMGQIRIRIM-----PTDVPOGLLCDLMSDPDKOVLGNGEND- 220

QY 333 RGAGWFGPDYTDNPLQHRRLSVIRSHCKRNGHSPMDNKIITTFASANTYVAGSNG 392
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 221 RGVSTFGAEVVAKFIAHKDLIDLCRAHQVVEDGHEFAKQDLVTLFSAPTYCGSEPDNAG 280
```

```
QY 393 AYIRLNQMLMPHFVQYISAASQTK 416
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 281 AMMSVDETLMCSF-QILKPAEKKK 303

RESULT 3
US-11-109-156-37
/ Sequence 37, Application US/11109156
/ Publication No. US20050250144A1
/ GENERAL INFORMATION:
/ APPLICANT: Toshio Isogai
/ APPLICANT: Takao Isogai
/ APPLICANT: Tetsuo Nishikawa
/ APPLICANT: Koichi Hayashi
/ APPLICANT: Kaoru Otsuka
/ APPLICANT: Jun-ichi Yamamoto
/ APPLICANT: Shizuko Ishii
/ APPLICANT: Tomoyasu Sugiyama
/ APPLICANT: Ai Wakamatsu
/ APPLICANT: Keiichi Nagai
/ APPLICANT: Tetsuji Otsuki
/ APPLICANT: Shin-ichi Fumahaashi
/ APPLICANT: Chiaki Senoo
/ APPLICANT: Jun-ichi Nezu
/ TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
/ FILE REFERENCE: 06501-099002
/ CURRENT APPLICATION NUMBER: US/11/109,156
/ CURRENT FILING DATE: 2005-04-19
/ PRIOR APPLICATION NUMBER: US/10/060,065
/ PRIOR FILING DATE: 2002-01-29
/ PRIOR APPLICATION NUMBER: PCT/JP00/05061
/ PRIOR FILING DATE: 2000-07-28
/ PRIOR APPLICATION NUMBER: US 60/159,590
/ PRIOR FILING DATE: 1999-10-18
/ PRIOR APPLICATION NUMBER: US 60/183,322
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: JP 11-248036
/ PRIOR FILING DATE: 1999-07-29
/ PRIOR APPLICATION NUMBER: JP 2000-118776
/ PRIOR FILING DATE: 2000-01-11
/ PRIOR APPLICATION NUMBER: JP 2000-183767
/ PRIOR FILING DATE: 2000-05-02
/ PRIOR APPLICATION NUMBER: JP 2000-241899
/ PRIOR FILING DATE: 2000-06-09
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 37
/ LENGTH: 323
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-11-109-156-37

Query Match          11.8%; Score 405.5; DB 7; Length 323;
Best Local Similarity 32.4%; Pred. No. 2.8e-24;
Matches 105; Conservative 58; Mismatches 126; Indels 35; Gaps 11;

QY 102 KKHIDLLIVFRKRGNR-----LHPKTV-ALLIREAKSKIKOLPNISPVSTAVSQOVT 154
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 6 KLNIDSIIORLLEVRGSKRGKVVQLOENIRIGLCLKSREIFLSQ-----PILBLEAPLK 60

QY 155 VCGDHLHGKLDLLVVLHKNGLPSSSNPVYFNGDFVDRGKRGLEVVLLLSLYLAPNNAV 214
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 61 ICGDHLHGQYDDLRLREFYGGPPPSN-YLFLGDYVDRGKOSLETICLLAYKIKYKPENF 119

QY 215 LNRGNHDSVMNARVGFIREVESKYPRNHKRIIAFIDERYRWPLGVSYVNSKVLVHGCF 274
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 120 LNRGNHSCASINRIYGFYDECKRRY--NIKLMKTFTD-CFNCPLPIAAYDEKIFCCHGGL 176

QY 275 S-DSTSLDLIKSIDRGKTVSILRPPLTDGEPIDKTEWQOIPDIMSDP-QATMGCVPTLT 332
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 177 SPDLQSMGQIRIRIM-----PTDVPOGLLCDLMSDPDKOVLGNGEND- 220
```

QY 333 RGAGWGPDPVTNDFLORHRLSYIRSHCEKPNHEFMHDKITITFSASNYTAISGNKG 392
DB 221 RGVSFTFGAIVAKFPLKHDLDLCRAHQVVDGVEFPARQVLTLFSAFNYGCEPFDNAG 280
QY 393 AYIRLNQMLPHFVQYISAASQTK 416
DB 281 AMMSVDETLWCSP-QILKPAEKCK 303
RESULT 4
US-11-109-156-34
; Sequence 34, Application US/11109156
; Publication No. US20050250144A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isogai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Makamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-ichi Funahashi
; APPLICANT: Jun-ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
; TITLE OF INVENTION: PHOSPHATASE
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/11/109,156
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: US/10/060,065
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-109-156-34
Query Match 11.4%; Score 391.5; DB 7; Length 521;
Best Local Similarity 31.1%; Pred. No. 6.8e-23; Indels 47; Gaps 10;
Matches 101; Conservative 56; Mismatches 119;
QY 102 KKHIDLLIDVFRKRGKRLHPKYVALILREAAKSLKQLPNISPVSTAVSQOVTVCGLDNG 161
DB 40 KPRVDILKALHMKR--GRLEBSVALRITREGASILRQEKVLDDIDA----PVTVCGLDNG 93
QY 162 KLDLVLVLAHKGKLPSSSNVYVRGDPVDRGKGLVLLLSLYLAFPAVLAFLNKGNH 221
DB 94 QFPLDKMLGFVGVGSPANTR-YLPLGVDYVDRGYFISICVLYLWALKITLYPRTLFLRGNH 152
QY 222 DSVNNAARYGFIRESVSKYPRNHRKRIILAFIDVYVRMLPLGSLVLSRVLIVHGFS-DSTSL 280
DB 153 CRHLTEFTFKQCKCKIKY---SERVYDCAKDAPDCPLAALMQQFLCVHGSLSPRLNTL 209

QY 281 DLKSIDRCKRYVSLIAPPLTDGEPPLDKTEWQQLFDITMSDPQATMG-----CVPTLR 333
DB 210 DDRIKDLRFK-----BPP-----AYGPMCDILMSDPLBDFGNEKTOBHPTNVTVR 254
QY 334 GAGWGPDPVTNDFLORHRLSYIRSHCEKPNHEFMHDK-----ITITFSASNYTAI 387
DB 255 GCSYFTSYPAVCEPFLQHNLLSLIRAHAEQADGIRYRTRRSQTTGFPSSLITITFSAPNYLDV 314
QY 388 GSNKGAIVIRLNQML-----PH 404
DB 315 YNNKAALVKYENNVNMIROFNCSPH 339
RESULT 5
US-11-109-156-24
; Sequence 24, Application US/11109156
; Publication No. US20050250144A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isogai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Makamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-ichi Funahashi
; APPLICANT: Jun-ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
; TITLE OF INVENTION: PHOSPHATASE
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/11/109,156
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: US/10/060,065
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-109-156-24
Query Match 11.0%; Score 379; DB 7; Length 309;
Best Local Similarity 31.1%; Pred. No. 2.9e-22; Indels 28; Gaps 7;
Matches 91; Conservative 62; Mismatches 112;
QY 119 KILHPKYVALILREAAKSLKQLPNISPVSTAVSQOVTVCGLDNGKLDLVLVLAHKGKLPSS 178
DB 22 QLESQVKSILCEKAKKILTKESNVQEVRC---PVTVCGLVHGQFHDLMELFRIGKSPD 77
QY 179 SNPYVNGDPVDRGKRGKGLVLLLSLYLAFPAVLAFLNKGNHDSVMAARYGFIRESVSK 238
DB 78 TN-YLPMGDYVDRGYSVETVLLVALKVRIRERTITLKGNHSSROI TOVYGFYDECLARK 136


```

QY 350 RHRLSVIRSHCK-PNGHFMHDK---ITTFASNYAIGSNKAYIRLNOLMPH- 404
DB 1213 BRKTF--KACANPBEPEKXMDKYTLTV-----SKVYIISIBIISTHS 1259
QY 405 -FVQYISAQTRLSFKQRMGIV-----ESSALXELAVMRD-----HRDELBD 448
DB 1260 ILLEHODALAPKNDLSELGLSGVPTVESFLGAVDPNDPNKANTLSQLSKTEISL 1319
QY 449 EF-RKYDPDQSGYISIMCKMMENTKGLPWRILDKLAPGDSQKXNNRTLDLDT 507
DB 1320 VLTSTKDIEDGALISR---SIMGTKKLIID--VIRNQ--PG-----VTLFEIILT 1364
QY 508 DYLEADGMSVMDALYANKASLVAFNIIADNNGEITLDBEFETAILDV-AHMRGAY 566
DB 1365 PATAOQEVHATDM---VSRA-----MIDSR---TPBMRKSGSMIEDAQLPLEQ 1408
QY 567 SKAEMLEKCRMDLNG---DGKVD--LNEPLAFLSDLHRKEQDE-----608
DB 1409 KKRKIQRNRTLEQCHVSESKYODILNEIAKDIPNQRIRYKRAELAKIQQTLNAIN 1468
QY 609 -----NIRRRST-----GRPSVAKTATDPVTLLADKISK 637
DB 1469 KKAAPYEQINYYDYIKTCLDNLKXKNTRSIKLDGKEBPQAKRA-KPVYTAAXLHE 1527
QY 638 NTLVVEHD 645
DB 1528 KGVLDID 1535

```

RESULT 8
US-10-995-561-810
Sequence 810, Application US/10995561
Publication No. US20050272054A1

GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 810
LENGTH: 548
TYPE: PRF
ORGANISM: Homo sapiens
US-10-995-561-810

Query Match 3.2%; Score 111; DB 6; Length 548;
Best Local Similarity 18.6%; Pred. No. 0.4;
Matches 86; Conservative 67; Mismatches 175; Indels 134; Gaps 19;

```

QY 273 GSDSTSLDLSIRGKXVSLRPP--LTDGEPDKTWMQOIFDIMSDDPQATMGCVN 330
DB 91 GPS-SFARSLVPSSE-LHLSLKSSKVTNGPEKNSSSPSSVDYAAAGPRKL----- 142
QY 331 TLRGAGWFGPVTNDFLQHRHLSVYIRSHCKPNG-----HEFMHDKII----- 376
DB 143 ---SSGALYGRPRSTFPNSHRFQVLVTSK--APSGDLIDKXSELSNNQLPPTPTLKT 196
QY 377 ---TFPSANYYAIGSNKAYIRLNQMLPHFVQYISAASQTRLSFKQRMGIVESSALK 433
DB 197 EAKSFLSQYRYTTPAKRKDPFD-----QRIEATQTE-LSFKSELGTAFRTKMT 245
QY 434 BLAVMRDRH-----DELEDEFKRY-----453
DB 246 DEBNIKQANSNCVTYAKKIAPLRLEGHDSITWDEIKDIALQHSPPRACQYSLKPPSTR 305
QY 454 ---DPKDSGYISISHCKMMENTKGLG-----PWRLLRDLKLAFTDSQKXNNRTLDL 504
DB 306 KIYSDBEELLYSFTB--DVTDEILKGLFSNRFLERLPERHIIKQKHLEBBERHLLHY 363

```

```

QY 505 LDPTVILEADGMSVMDALYANKASLVAFNIIADNNGEITLDBEFETAILDVAMRG 564
DB 364 LKVDAGCTSEENSVQND-----VDMLVNFPDEKAGSEBPELKNSEVTTIQOERO 414
QY 565 AYSKAEMLKCRMDLNGDKVDLNE-----PLEAFRLSLHR--KEQDENIRRR 613
DB 415 QYOKA-----LDMLSAPRDEBEIPSPTEFPMPYIKSGHSEGVIIQQVNDFTNLT 466
QY 614 ST---GRPSVAKTATDPVTLLADKISKNTLVVEHDIDPTDC 652
DB 467 STLDENHPSISDITDRET-----SVNVIEGSDPEKVE 500

```

RESULT 9
US-10-995-561-813
Sequence 813, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 813
LENGTH: 567
TYPE: PRF
ORGANISM: Homo sapiens
US-10-995-561-813

Query Match 3.2%; Score 111; DB 6; Length 567;
Best Local Similarity 18.6%; Pred. No. 0.42;
Matches 86; Conservative 67; Mismatches 175; Indels 134; Gaps 19;

```

QY 273 GSDSTSLDLSIRGKXVSLRPP--LTDGEPDKTWMQOIFDIMSDDPQATMGCVN 330
DB 110 GPS-SFARSLVPSSE-LHLSLKSSKVTNGPEKNSSSPSSVDYAAAGPRKL----- 161
QY 331 TLRGAGWFGPVTNDFLQHRHLSVYIRSHCKPNG-----HEFMHDKII----- 376
DB 162 ---SSGALYGRPRSTFPNSHRFQVLVTSK--APSGDLIDKXSELSNNQLPPTPTLKT 215
QY 377 ---TFPSANYYAIGSNKAYIRLNQMLPHFVQYISAASQTRLSFKQRMGIVESSALK 433
DB 216 EAKSFLSQYRYTTPAKRKDPFD-----QRIEATQTE-LSFKSELGTAFRTKMT 264
QY 434 BLAVMRDRH-----DELEDEFKRY-----453
DB 265 DEBNIKQANSNCVTYADAKKIAPLRLEGHDSITWDEIKDIALQHSPPRACQYSLKPPSTR 324
QY 454 ---DPKDSGYISISHCKMMENTKGLG-----PWRLLRDLKLAFTDSQKXNNRTLDL 504
DB 325 KIYSDBEELLYSFTB--DVTDEILKGLFSNRFLERLPERHIIKQKHLEBBERHLLHY 382
QY 505 LDPTVILEADGMSVMDALYANKASLVAFNIIADNNGEITLDBEFETAILDVAMRG 564
DB 383 LKVDAGCTSEENSVQND-----VDMLVNFPDEKAGSEBPELKNSEVTTIQOERO 433
QY 565 AYSKAEMLKCRMDLNGDKVDLNE-----PLEAFRLSLHR--KEQDENIRRR 613
DB 434 QYOKA-----LDMLSAPRDEBEIPSPTEFPMPYIKSGHSEGVIIQQVNDFTNLT 485
QY 614 ST---GRPSVAKTATDPVTLLADKISKNTLVVEHDIDPTDC 652
DB 486 STLDENHPSISDITDRET-----SVNVIEGSDPEKVE 519

```

RESULT 10
US-10-995-561-812
Sequence 812, Application US/10995561

```
/ Publication No. US20050272054A1
/ GENERAL INFORMATION:
/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
/ TITLE OF INVENTION: DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001559
/ CURRENT APPLICATION NUMBER: US/10/995,561
/ CURRENT FILING DATE: 2004-11-24
/ NUMBER OF SEQ ID NOS: 85702
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 812
/ LENGTH: 599
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-10-995-561-812

Query Match          3.2%; Score 111; DB 6; Length 599;
Best Local Similarity 18.6%; Pred. No. 0.46;
Matches 86; Conservative 67; Mismatches 175; Indels 134; Gaps 19;

QY 273 GSDSTSLDLISIRGKVSILRPP--LTDEPLDKTWQOIFDIMSDDPATMGCVPN 330
DB 142 GFS-SPASLVPSSER-LHLSLHKSRYITNGPERKSSSPSSVYAAAGPRKL----- 193
QY 331 TLRGAGVWPGPDVTDNFIQRHRLSYIRSHCKPNG-----HEFMHDKLI----- 376
DB 134 ---SSGALYGRPRSTGTFPSHRFOLYISK--APSGDLIDKISELFSNKQLPFTPTTKT 247
QY 377 ---TFPSANVYAIGSNKGAIVIRLNNQMLPHFVOYISAASQTKRLSPKQRMGIVESALK 433
DB 248 EAKSEFLSQRYRYTPARKKQFTD-----QRIEATQTE-LSPKSELGTAATKQMT 296
QY 434 ELAVMRDR-----DELEDFRKY----- 453
DB 297 DSENMKIKQASNCVYDAKEKIALPLBGHDSITWDEIKDQALQHSPPRAMCOYSLKPPSTR 356
QY 454 ---DPKDSGYISISHMKCMENVTTLG-----PMRLRLDLAPQDSOKVNYRTLDL 504
DB 357 KLYSDEBELLYISFIE--DYTDILKGLFSNRFLERIRIKONKHLSEKMHLLHV 414
QY 505 LDTVILEADGMSVMDALYANKASLVAIFNIIDANSGETLDEFEFTALDILVAHMPG 564
DB 415 LKVDLGGCTSEENSVKQND-----VDMLVNPFDEKAGNBPRLKNSSVYTIQDSRQ 465
QY 555 AYSKAEMLKCMMDLNGGKVDLNE-----FLAEFLSLDLR--KEQDENIRRR 613
DB 466 QYOKA-----LDMLSAPDSENBIFPSPTEFFMPIYKSRGSGVLIIOQVNDSTNLET 517
QY 614 ST---GRPSVAKTATDPVTLLADKISKNTLVVEHIDPDC 652
DB 518 STLDENHPSI9DSLTDRET-----SVNVIRGSDPEKVR 551

RESULT 11
US-11-074-176-158
/ Sequence 158, Application US/11074176
/ Publication No. US20050250135A1
/ GENERAL INFORMATION:
/ APPLICANT: Kluenhammer, Todd R.
/ APPLICANT: Russell, William M.
/ APPLICANT: Altermann, Eric
/ APPLICANT: McAniff, Olivia
/ APPLICANT: Perill, Andrea Azcarate
/ TITLE OF INVENTION: Nucleic Acid Sequences Encoding
/ TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
/ FILE REFERENCE: 5051-694
/ CURRENT APPLICATION NUMBER: US/11/074,176
/ CURRENT FILING DATE: 2005-03-07
/ PRIOR APPLICATION NUMBER: 60/551,161
/ NUMBER OF SEQ ID NOS: 381
/ SOFTWARE: FastSeq for Windows Version 4.0
```

```
/ SEQ ID NO 158
/ LENGTH: 709
/ TYPE: PRF
/ ORGANISM: Lactobacillus acidophilus
US-11-074-176-158

Query Match          3.1%; Score 108; DB 7; Length 709;
Best Local Similarity 17.4%; Pred. No. 1;
Matches 113; Conservative 97; Mismatches 176; Indels 264; Gaps 31;

QY 83 DLVESFGDINVAKIELPI-RKNHIDLIDVF-RKKRGNRLHPKVALILREAAKSLQLP 140
DB 68 DLVER---AKNNKFPDVGIDQIDNVILSRKKNPV----- 104
QY 141 NISPVSTAVSOQV-----TCGDLHGGLDGL-LVHLKNGLPSNSNRYVNGDFVDRCK 193
DB 105 LIGPAGVGTISIVGGLAERIAAGNVPAMVANNHIIISVINIMVAGSS---LRGSEERLK 161
QY 194 RGLBVLTLTLILYLAFPAVAFILNRGNHDSVNNARYGFIREVESKYPRNHKRIILFIDV 253
DB 162 KYID----- 180
QY 254 YRMPLGSLVINSRYLIVHGFSDDSTSLDLISIRGKVSILRPLTDGE-----PLDKT 308
DB 181 HN-----LVGAGSTDSN-----NNGDAANILKPLAAGELKIGATTTS 220
QY 309 EMQOIFDIMSDDP-----QATMGCVPTTLGAGVWFGPDVTDNFIQRHRLSYIRSHCK 362
DB 221 EFORI---EKDPALSRPRQAVOYEPET-----DVAIKLIEGLKKYEDYHHV- 265
QY 363 KPNGHFHFMDNKITTFPSANVYAIGSNKGAIVIRLNNQMLPHFVOYISAASQTKRL----- 418
DB 266 ---KTYDSSLKLAVERISRYI---QGRY-----LPDKAIDMLBAGAKKALLVOP 309
QY 419 ---SFKQRMGIVESALK-----ELAVMRDRDELDEDFRKYDPKDSGYISISHMC 467
DB 310 TDEKSLKQIISALBKAKBAKABEDYDCAAEIKKIALBSGLKSVDSKPEVATKOIY 369
QY 468 KYMERVTLGLPMLRLKRLAPGDSOK-----VNYRTLDLDTVILEA----- 514
DB 370 QIIBOKTIPMS-BLHADBQKNLDLAKKLKKNVLDODRAIDVI-TDAIARQIIPKSDR 427
QY 515 -----ADGMSVMDALYANKASLVAIFNIIDANSGETLDEP-ET 554
DB 428 PTGSELTLPYGVGTETLAKQLAIO--LFGKKEHLI-----RLDMSEYQDEMA 473
QY 555 IDLVAHMPG-----AYS-----KA-----EMLEKCRMDL 580
DB 474 VNKLIGSAPGYVYGEGQLTEKVRHQPYSLILFDEIRKAMPQYENALLQIMDDGRLTDA 533
QY 581 NG-----DQKVDLNEFLA---FRUSDILR 602
DB 534 QGRVSPFDITLIMTSNAGPSDKLLEDKVDQDKLISALBNYFRPEFLNR 583

RESULT 12
US-11-202-516-10
/ Sequence 10, Application US/11202516
/ Publication No. US2006008465A1
/ GENERAL INFORMATION:
/ APPLICANT: Steinna, Lucilla
/ APPLICANT: Mouritsen, Soren
/ APPLICANT: Gautam, Anand
/ APPLICANT: Dalum, Iben
/ APPLICANT: Haaning, Jesper
/ APPLICANT: Leach, Dana
/ APPLICANT: Nielsen, Klaus
/ APPLICANT: Karlsson, Gunilla
/ APPLICANT: Rasmussen, Peter
/ TITLE OF INVENTION: Novel Methods for Therapeutic Vaccination
/ FILE REFERENCE: 4614-0107PUS2
/ CURRENT APPLICATION NUMBER: US/11/202,516
/ CURRENT FILING DATE: 2005-08-11
```

```
; PRIOR APPLICATION NUMBER: US 09/806,703
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: PCT/DK99/00525
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: DK 1998 01261
; PRIOR FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: US 60/105,011
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-202-516-10

Query Match      3.1%; Score 107.5; DB 7; Length 694;
Best Local Similarity 19.9%; Pred. No. 1.1;
Matches 109; Conservative 72; Mismatches 194; Indels 173; Gaps 29;

QY      38 EVASEDDQAEIYKFENDLIKHPQAGKQOYGSAAHVSLDDKDLVEEFG-DIVNAKI 96
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      4 EFLHELKANKIKKFLNPFTR-TPHLAGTQNNBELAKQI-----HQQWKEFGDLV- 52
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      97 ELPFRKNHIDLIDVFRKKRGNRLHPKYVALLIREAAK-----SLKQLP-----NIS---P 144
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      53 EL-----SHYDVLLSY-----PNKTHPNYISIIINEDGNEIFKTSLSHQPPGYENISDVVP 103
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      145 VSTAVSQOYVCGDL-----HGKLDLVLVHLKNGLPSSNPVYFNGDFVDRGRGGLVLL 200
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      104 PYSAPSPQGTPEBDLVYVYARTEDFPKLEREMKISCSGKIYIARYGKFRGN----- 156
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      201 LLLSLYAFNNAVFLNRGNHEDSVNNARYGFTREVESKYPRNHKRIILAFIDEVYRW-LPL 259
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      157 MYKNAQLAGAKGMIL-----YSDP---ADY-FVPAVKS-YPDG-----WNLPG 194
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      260 GSYLNSRVLLVHGFGSDSTSLDLIKSIDRGKYVSIIRPPLTDSGPDLKTEWQO----- 312
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      195 GGVRGNVNLNGA-----GDLPLTGYPANBHAAYRHELTNAV 232
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      313 -----IFDIMSPDQATMGCVPTLNRGAVW-----FGPDVTDNFLQHRRLSYVIR 358
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      233 LPSIPVHPITGYDDAQKLEHMGAPAPDSSWKGGLKVPYVNGGFAGNF-----S 252
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      359 SHECKENGHEFMDNKITITFASANYAIGSNKGAAYIRLNQMLPHFVOYISAASQTKRL 418
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      283 TOKVKNHHSYTKYTRI-----YVIGTLKGA-----LEPD--RYVILGHRDAM 325
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      419 SF-----KQRMGIYESSALKELAVRMRDHDELEDFRKYDPPDSGYISISHKCYMENY 473
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      326 VFGIDPQSGAAVVEIYVRSFGTLKKKGRPRRTIIFASWDABEFGILGSTEWABEHS-- 383
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      474 TKGLPFWLLR-----DKLAPGDSQKN-----YNRRLDLDLDTVILLEAA 515
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      384 -----RLQERGVAYINADSSIEGNYTLRVDCPLMTSLVYNLTKELQSPDEGFE--- 433
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      516 DQMSVMDA 523
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      434 -GKSLYDS 440
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-11-202-516-8
; Sequence 8, Application US/11202516
; Publication No. US20060008465A1
; GENERAL INFORMATION:
; APPLICANT: Steinaa, Lucilla
; APPLICANT: Mouritsen, Soren
; APPLICANT: Gautam, Anand
; APPLICANT: Dalum, Iben
; APPLICANT: Haaning, Jesper
; APPLICANT: Leach, Dana
; APPLICANT: Nielsen, Klaus
```

```
; APPLICANT: Karlsson, Gunilla
; APPLICANT: Rasmussen, Peter
; TITLE OF INVENTION: Novel Methods for Therapeutic Vaccination
; FILE REFERENCE: 4614-0107PUS2
; CURRENT APPLICATION NUMBER: US/11/202,516
; CURRENT FILING DATE: 2005-08-11
; PRIOR APPLICATION NUMBER: US 09/806,703
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: PCT/DK99/00525
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: DK 1998 01261
; PRIOR FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: US 60/105,011
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 752
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-202-516-8

Query Match      3.1%; Score 107.5; DB 7; Length 752;
Best Local Similarity 19.9%; Pred. No. 1.2;
Matches 109; Conservative 72; Mismatches 194; Indels 173; Gaps 29;

QY      38 EVASEDDQAEIYKFENDLIKHPQAGKQOYGSAAHVSLDDKDLVEEFG-DIVNAKI 96
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      62 EFLHELKANKIKKFLNPFTR-TPHLAGTQNNBELAKQI-----HQQWKEFGDLV- 110
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      97 ELPFRKNHIDLIDVFRKKRGNRLHPKYVALLIREAAK-----SLKQLP-----NIS---P 144
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      111 EL-----SHYDVLLSY-----PNKTHPNYISIIINEDGNEIFKTSLSHQPPGYENISDVVP 161
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      145 VSTAVSQOYVCGDL-----HGKLDLVLVHLKNGLPSSNPVYFNGDFVDRGRGGLVLL 200
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      162 PYSAPSPQGTPEBDLVYVYARTEDFPKLEREMKISCSGKIYIARYGKFRGN----- 214
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      201 LLLSLYAFNNAVFLNRGNHEDSVNNARYGFTREVESKYPRNHKRIILAFIDEVYRW-LPL 259
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      215 MYKNAQLAGAKGMIL-----YSDP---ADY-FVPAVKS-YPDG-----WNLPG 252
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      260 GSYLNSRVLLVHGFGSDSTSLDLIKSIDRGKYVSIIRPPLTDSGPDLKTEWQO----- 312
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      253 GGVRGNVNLNGA-----GDLPLTGYPANBHAAYRHELTNAV 290
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      313 -----IFDIMSPDQATMGCVPTLNRGAVW-----FGPDVTDNFLQHRRLSYVIR 358
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      291 LPSIPVHPITGYDDAQKLEHMGAPAPDSSWKGGLKVPYVNGGFAGNF-----S 340
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      359 SHECKENGHEFMDNKITITFASANYAIGSNKGAAYIRLNQMLPHFVOYISAASQTKRL 418
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      341 TOKVKNHHSYTKYTRI-----YVIGTLKGA-----LEPD--RYVILGHRDAM 383
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      419 SF-----KQRMGIYESSALKELAVRMRDHDELEDFRKYDPPDSGYISISHKCYMENY 473
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      384 VFGIDPQSGAAVVEIYVRSFGTLKKKGRPRRTIIFASWDABEFGILGSTEWABEHS-- 441
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      474 TKGLPFWLLR-----DKLAPGDSQKN-----YNRRLDLDLDTVILLEAA 515
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      442 -----RLQERGVAYINADSSIEGNYTLRVDCPLMTSLVYNLTKELQSPDEGFE--- 491
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      516 DQMSVMDA 523
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      492 -GKSLYDS 498
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-10-793-626-2964
; Sequence 2964, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
```

/ TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 / FILE REFERENCE: P03480US
 / CURRENT APPLICATION NUMBER: US/10/793,626
 / CURRENT FILING DATE: 2004-03-04
 / PRIOR APPLICATION NUMBER: 60/164,258
 / PRIOR FILING DATE: 1999-11-09
 / NUMBER OF SEQ ID NOS: 4472
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 2964
 / LENGTH: 5024
 / TYPE: PRT
 / ORGANISM: Artificial Sequence
 / FEATURE:
 / OTHER INFORMATION: Description of Artificial Sequence: synthetic
 / OTHER INFORMATION: amino acid sequence
 / FEATURE:
 / NAME/KEY: MOD_RES
 / LOCATION: (5024)
 / OTHER INFORMATION: variable amino acid
 / US-10-793-626-2964

Query Match 3.1%; Score 105; DB 6; Length 5024;

Best Local Similarity 20.0%; Pred. No. 30;

Matches 149; Conservative 105; Mismatches 274; Indels 218; Gaps 36;

QY 53 NDLIKHM-----PQAGRKNQYGSAAHVSVDLDDKDLVREPGDIYNAKIS 97
 DB 3605 NDEIKEMNTLANNAOKSISDMISHALRTREYKQLDQAATLDEAMSLSDKQVITDTT 3664
 QY 98 LP-----IRKNHIDLID---VERKRGKRLHPRVYALLREAAKSLKQLPNISPV 146
 DB 3665 LPVYTESBDEKKEKQVTSQAQALIDKINGSVNSLDQVRQALQQLTQASBNLDGQRYE 3724
 QY 147 TAVSQQVTCGDLHGK---LDDLVLVLRKNGLPSSSPYFVNGDFVDRGRLGLVLLILLS 204
 DB 3725 EA-----KXHANQITDQL---THLSLQOQT-----AESVYNAATKLEE 3760
 QY 205 LYLAFPAVFLNRG-----NHSDVYNNARYGFRIVESKYPHKKRIILAFIDEVYRW 256
 DB 3761 IATASNALALANKVMKLEQFINHADSISNS-----DNTRQADDKILIAVDLBLENG 3812
 QY 257 -----LPLGSVLSRVLIIVHGCF-----SDSTSLDLTKSIDR--GKVVSTL 295
 DB 3813 ODIOKSNATONBEAKQALQOLINAEISL--NGFERLHARPRALVEYIKLEKINNQKSL 3870
 QY 296 RPPLTGGEPLDK-----TEWQQLFDIM-----NSDPQATMGCVP----- 329
 DB 3871 EDKVTQSHDLLEHLVNSGTMLNDIMGELANAIYNNYAPTKASINYINADMLRKDNFTQ 3930
 QY 330 -----NTLGGAGVWFGPQVY---DNFLOHRHLSYVIRSHCEKPNHGFHMDKTI- 376
 DB 3931 AINNADALANKTQGMIDFRALIDTFKODIFKTYDALNGIERLTAASKAKELIDSLKPIFN 3990
 QY 377 -TTPSASNYVAIGSNKAYI-----RLNNQMLPHFVQ-----YISAA 412
 DB 3991 KQGFTHANDEINVTNSIAQISRLVNOAFPLDAMKSLRDELINNAQAPF--VQSSSVNTIND 4048
 QY 413 SQTKLSLFGQRMKIVSSSALKELAVMRDRDELDEFKRY--DPRD--SGYISISH-W 466
 DB 4049 E-----DLKQGFPHALSNAKVLAKENKGLDEIQLEGKQVIEDTKDLMLNGIORSKAK 4103
 QY 467 CKYMNVTYKLGLPWRLRLDLKAPGTSOK-----VNNRRTLDLLDTVIIAABADGWSVW- 521
 DB 4104 AKRIQVVOGLSYI-----NDAQRHIAESNTHNSDDLSLANTLSKASDLDNAMK 4152
 QY 522 ---DALYANKASIVAFNIIDANSGEITLDEF---ETAIDLVAHMPGAYSKAEMLERK 575
 DB 4153 DLRLDTLESNSTSVNSVYINADKNLQIEPDEALQQAASATSSKTSNPAITIEVGLSLQA 4212
 QY 576 RMMU-----LNGDQKV--DLNEPLFAFR--LSDLRKQCOQDENIRRSTGSPVAKTYT--- 625
 DB 4213 -YDTYONALNGEORLATERKSKDKLTKGLDLNKAQLEB-----VTNNVNSAMTILTEIS 4265

QY 626 -----DPTVLADKISKNTLV 641
 DB 4266 QLTOSTLKNKMKLLRDKL--KTLV 4289

RESULT 15

US/11/062

/ Sequence 4, Application US/11062471A

/ Publication No. US20050255093A1

/ GENERAL INFORMATION:

/ APPLICANT: SHORE, Clifford Charles

/ APPLICANT: SUTTON, John Mark

/ APPLICANT: HALLIS, Bassem

/ APPLICANT: SILMAN, Nigel

/ TITLE OF INVENTION: Delivery of Superoxide Dismutase to Neuronal Cells

/ FILE REFERENCE: 1581.0800001

/ CURRENT APPLICATION NUMBER: US/11/062,471A

/ CURRENT FILING DATE: 2005-02-22

/ PRIOR APPLICATION NUMBER: 09/831,050

/ PRIOR FILING DATE: 1999-11-05

/ PRIOR APPLICATION NUMBER: PCT/GB99/03699

/ PRIOR FILING DATE: 1999-11-05

/ PRIOR APPLICATION NUMBER: GB 9824282.9

/ PRIOR FILING DATE: 1998-11-05

/ NUMBER OF SEQ ID NOS: 11

/ SOFTWARE: PatentIn Ver. 2.1

/ SEQ ID NO 4

/ LENGTH: 1070

/ TYPE: PRT

/ ORGANISM: Artificial Sequence

/ FEATURE:

/ OTHER INFORMATION: Construct comprising Mn-SOD from B. steaerothermophilus, a linker,

US/11/062,471A-4

Query Match 3.0%; Score 104.5; DB 7; Length 1070;

Best Local Similarity 18.7%; Pred. No. 3.4;

Matches 116; Conservative 98; Mismatches 182; Indels 223; Gaps 33;

QY 10 IFIOKRY-----RRHQARERMRRCMQIIFONLE-----YASEQOABLYKRFND-- 54
 DB 478 LIVAQMISTVNTQFTTIEGMYKALNTYQ-AQLSEIIRYRYNYSEKKSNTINDFNIN 536
 QY 55 -LIGMPQAGRKQYOGSAHVSVDLDDK--DVLVEFGDIYNAKIEPIRKHIDLID 110
 DB 537 SKLNGCINQALDININFPNGCSVSLMKMPLAVKLLDPNT-----LKNLLANYIDE 591
 QY 111 VERKKRGRLHPRKYVALILREAAKSLKQLPNISPVSTAVSQQVTCGDLHGKTL----- 163
 DB 592 -----NKLV--LIGSAVEYKSKVKNYKLTIMPDLISITYTDTILIEFKNYNSIILNN 642
 QY 164 -----DOLLVLVLRNG-----LPSGSPN-----YFNGDF 188
 DB 643 IILNLRVNDNMLIDISGYGAKVEYVYDVELNDKQFKLTSSANSKIRVTQONQIIFNSVF 702
 QY 189 VDRGRGLLEVLILLSLYLAPF--NAVFLNRGNHSDVMNA----- 227
 DB 703 LD-----FVSIFPIRIKPYKQNDGILQNYIHIEYITIIINCKNNNSGKISIRGNRIIW 752
 QY 228 -----RYGPIREVESKYPRNHKRIILAFIDEVYRW--LPLGSVLSRVLIIVH 271
 DB 753 TLIDINGKTKSVFPEYN--IREDISY-----INRWFPVTITNNLNNAKIYIN 798
 QY 272 GGFPSSTSL-DLKSIDGKYVSIIRPPLTDEPLDKTR--WQQLFDIMSDPQATMGCV 328
 DB 799 GLRLSNTIYKDIRVIANGELIFKL-----DGD-IDRQOIFIMKXYFSI----- 840
 QY 329 PTLRGGAGVWFGPQVDTNDFLOHRHLSYVIRSHCEKPNHGFHMD-----NKLITIFS 380
 DB 841 -----FTRELSSQNSIER--YKIQSIS-----EYAKDWGMPPLMINKETVYFVN 881
 QY 381 ASNYVAIGSNKAYIRLNNQMLPHFVOYISAASQTKLSLFGQRMKIVSSSALKELAVNR 440
 DB 882 A-----GNKSYIKLKDKSPVGEILTRRSKYNQSKYINRYDLYIGF-----KFLIRRK 929


```
QY      441 DHDHDELEDER-RKYDPROGSGYSISHWCKMENVNFKGLPRLLEDK-----LAP 489
           ||| |
Db       930 SNGQSINDIVRKED-----YYLDPF-----NNQBRVRYTYKKFKBEKBLFAP 976
           ||| |
QY      490 GTDSQKVNRRTIIDLDDTD 508
           ||| : ||| : ||| :
Db       977 ISDSDEF-YN-TIQIKED 993
```

RESULT 16
US/11/062

```

Sequence 7, Application US/11062471A
Publication No. US20050255093A1
GENERAL INFORMATION:
APPLICANT: SHONE, Clifford Charles
APPLICANT: SUTTON, John Mark
APPLICANT: HALLIS, Bassam
APPLICANT: SILMAN, Nigel
TITLE OF INVENTION: Delivery of Superoxide Dismutase to Neuronal Cells
FILE REFERENCE: 1581.0800001
CURRENT APPLICATION NUMBER: US/11/062,471A
CURRENT FILING DATE: 2005-02-22
PRIOR APPLICATION NUMBER: 09/831,050
PRIOR FILING DATE: 1999-11-05
PRIOR APPLICATION NUMBER: PCT/GB99/03699
PRIOR FILING DATE: 1999-11-05
PRIOR APPLICATION NUMBER: GB 9824282.9
PRIOR FILING DATE: 1998-11-05
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 1095
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Construct comprising a mitochondrial leader sequence from Human N
US/11/062,471A-7

```

| | | | | |
|-----------------------|--------------|----------------|-----------------|--------------|
| Query Match | 3.0% | Score 104.5; | DB 7; | Length 1095; |
| Best Local Similarity | 18.7% | Pred. No. 3.6; | | |
| Matches 116; | Conservative | 98; | Mismatches 182; | Indels 223; |
| | | | | Gaps 33; |

```

Qy 10 IFIOKWY-----RRHQARREKQRCNMQIIFONLS-----YASEQDAELYEPND-- 54
Db 503 LIVQOMISTVNTQTFYTTIKGMYKALNTQYA-QALBETIIRYRNYISEKKSININIDFNDIN 561
Qy 55 --LIKEMPOAGRKNOYOGSAHVSGLVDK--DDLVEBFGDLYVNAKIEBPIRRNHIDLID 110
Db 562 SKLNEGINDQALDININNPINGCSVSGLMKGMPLAVEKILDDPNT-----LKNULNTAYDE 616
Qy 111 VFRKKGRRLHPKYVALILREAAKSLKQJPNISPVSTAVSOOVTVCQDHLGKL----- 163
Db 617 -----NKLY--LIGSAEYKSKYKTKLKTMTDPDLSIYNTDITLIMFKYNSILANT 667
Qy 164 -----DDLVLVHLKNG-----LPSSNP-----YFNGDF 188
Db 668 IILNIIRYKDNMTLIDSGYGALEYVYDVELNDKQOFKLTSSANSKIRVTQNONITFNSVF 727
Qy 189 VDRGRGRGLVALLLSLTLAP--NAVFLRGNHDSVMA----- 227
Db 728 LD-----FVSFPMIRLIPKXKNDGIOYIHNEYITINCKMKNNSGWKISIRGNRIIW 777
Qy 228 -----RYGEIRFEVSEKYPNNHXRILLAFIDEVYRW--LPLGSVYANSRVLIVH 271
Db 778 TLIDINGKTSVPEBYN-IRREDISY-----IRMPFVITTNLNNAKIYIN 823
Qy 272 GGFSDSTSL-DLKSIDRGKYVSLRPPLTGEPLDKTE--WQOIFDIMSDDPOATMGCV 328
Db 824 GKLSNSTDKIDIREVIANGRIIFKL-----DQD-IDRQOIFWKKYFSI----- 865
Qy 329 PNTLRGAGVWGGPVDNTNPLQHRRLSYIRSHCEKPNHGFHMD-----NKITTFIS 380

```

```

Db      866  -----FNTLSQSNIIEER---YKIDSY-----BYLKDFWGNPLMYNKKRYEMEN 906

Qy      381  ASNTYTAIGSKNGAVIRLANNOLMPEFVOTISAASOTKELSFQRMGIVYESSALKELAVMR 440
Db      907  A-----GKNSTYIKLKDSPVGEILTRSKYQNSKYNRYRDYIGR-----KFILRRK 954

Qy      441  DHRDELEDEF-RKCDPRKDSGYTISIMCKMENTYKLGIPRYRLLRDK-----LAP 489
Db      955  SNSQSIINDPIVRKED-----YIYLDPF-----NLNQEWRYVYTKYFKKEBEKFLAP 1001

Qy      490  GTDSQKVNYNRRLDLDPTD 508
Db      1002  ISDSDEF-YN-TIQIKEYD 1018

```

RESULT 17

```

US-10-118-590-18
; Sequence 18, Application US/10118590
; Publication No. US20050277761A1
; GENERAL INFORMATION:
; APPLICANT: KENNETH RHODES, MARIA BETTY, HUI-PING LING, AND FRANK AN
; TITLE OR INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: NMT-070
; CURRENT APPLICATION NUMBER: US/10/118,590
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US/09/298,731
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-118-590-18

```

| | | | | |
|-----------------------|------------------|-----------------|--------|-------------|
| Query Match | 3.0%; | Score 103; | DB 6; | Length 270; |
| Best Local Similarity | 25.7%; | Pred. No. 0.59; | | |
| Matches 59; | Conservative 25; | Mismatches 58; | Indels | |

| | | | | |
|----|--|-----|--|-----|
| Qy | | 439 | MEDRR-----DELEDERKRTKDPSDGVISLSHCKWMENTKGLPRLRLARDKLA---- | 488 |
| Dd | | 64 | LRPARRPLDPDSVEDEF-----ELSTVCHRBEJGLOEQBQTFTTRRELDVLTYR | 112 |
| Qy | | 489 | -----PGTDSQKNTNRYTL-DLDTFD-----VILEAEADGMS | 519 |
| Dd | | 113 | GFKRCESGIVNEENPKQIYSQFPQGDSS--NYATFLFNAFDTHDSDSYSEDDVAGLS | 170 |
| Qy | | 520 | VM-----DALYANKASLIWALFNITIIDADNSGEITLDEFEFAIDLL--VAHMPGAYSKAEM | 571 |
| Dd | | 171 | VILRGTTIDRL--VMA-----FNLVDINKDCITKEBM--LDIMKSIFYDMWGKTYPAL | 220 |
| Qy | | 572 | LEKC-----RMMDNGDGKTKDLNFLEAFRLSLDLHRKEQODENIR | 612 |
| Dd | | 221 | RERAPREHVBSFFQKMDRNKGAVVTIEEPISSC-----QQDENIR | 261 |

RESULT 18

US-10-055-877-140
 : Sequence 140, Applicant US/10055877
 : Publication No. US20050288241A1
 : GENERAL INFORMATION:
 :
 : APPLICANT: Decisofaro, Marc
 : APPLICANT: Padisgar, Mutallidhara
 : APPLICANT: Miller, Charles
 : APPLICANT: Tcherenev, Velizar
 : APPLICANT: Zhong, Mei
 : APPLICANT: Anderson, David
 : APPLICANT: Ballinger, Robert
 : APPLICANT: Gerlach, Valerie
 : APPLICANT: Spytek, Kimberly
 : APPLICANT: Ratelli, Luca
 : APPLICANT: Kerkuda, Rameesh
 : APPLICANT: Guo, Xiaojia


```

; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 268
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-268

Query Match
Best Local Similarity 28.1%; Pred. No. 0.3;
Matches 39; Conservative 20; Mismatches 58; Indels 22; Gaps 6;

QY 483 LRDKLAPGTSQKVNRTLDLDTVILEADGM-----SYMDALYAN--KASLVA 533
DB 1 MADQL---TEBQIAEFKAPSLFDK-----GDGTTTKELGTVMSLQNPTEARLQD 51

QY 534 IFNIIDADNSGEITLDEFEFTAIDLVAHMPGAVSKAMLEKCRMDLNGDKVDLNEFLE 593
DB 52 MINEVDADNGGTTIDPEPFLT---MMARKKQDTSSEBIRAFVFDKDGNGYISAELRH 108

QY 594 AF-RLSDDLHRKEQDENIR 611
DB 109 VMTNLGKJLTDEEVDEMIR 127

RESULT 21
US-11-124-368A-269
; Sequence 269, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: C1001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; PRIOR FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 269
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-269

Query Match
Best Local Similarity 3.0%; Score 102; DB 7; Length 149;
Matches 39; Conservative 20; Mismatches 58; Indels 22; Gaps 6;

QY 483 LRDKLAPGTSQKVNRTLDLDTVILEADGM-----SYMDALYAN--KASLVA 533
DB 1 MADQL---TEBQIAEFKAPSLFDK-----GDGTTTKELGTVMSLQNPTEARLQD 51

QY 534 IFNIIDADNSGEITLDEFEFTAIDLVAHMPGAVSKAMLEKCRMDLNGDKVDLNEFLE 593
DB 52 MINEVDADNGGTTIDPEPFLT---MMARKKQDTSSEBIRAFVFDKDGNGYISAELRH 108

QY 594 AF-RLSDDLHRKEQDENIR 611
DB 109 VMTNLGKJLTDEEVDEMIR 127

RESULT 22
US-10-821-234-935
; Sequence 935, Application US/10821234
; Publication No. US20050255114A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmant, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 935
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-935

Query Match
Best Local Similarity 3.0%; Score 102; DB 6; Length 171;
Matches 39; Conservative 20; Mismatches 58; Indels 22; Gaps 6;

QY 483 LRDKLAPGTSQKVNRTLDLDTVILEADGM-----SYMDALYAN--KASLVA 533
DB 23 MADQL---TEBQIAEFKAPSLFDK-----GDGTTTKELGTVMSLQNPTEARLQD 73

QY 534 IFNIIDADNSGEITLDEFEFTAIDLVAHMPGAVSKAMLEKCRMDLNGDKVDLNEFLE 593
DB 74 MINEVDADNGGTTIDPEPFLT---MMARKKQDTSSEBIRAFVFDKDGNGYISAELRH 130

QY 594 AF-RLSDDLHRKEQDENIR 611
DB 131 VMTNLGKJLTDEEVDEMIR 149

RESULT 23
US-11-065-943-46
; Sequence 46, Application US/11065943
; Publication No. US20050250131A1
; GENERAL INFORMATION:
; APPLICANT: JEBSTIN, JEAN-LUC
; APPLICANT: VICHIER-GUERRE, SOPHIE
; APPLICANT: FERRIS, STEPHANE
; TITLE OF INVENTION: METHODS FOR OBTAINING THERMOSTABLE ENZYMES, DNA POLYMERASE I
; TITLE OF INVENTION: VARIANTS FROM THERMUS AQUATICUS HAVING NEW CATALYTIC ACTIVITIES
; FILE REFERENCE: 266426USOKIP
; CURRENT APPLICATION NUMBER: US/11/065,943
; PRIOR FILING DATE: 2005-02-25
; PRIOR APPLICATION NUMBER: US 10/787,219
; PRIOR FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 46
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-065-943-46

Query Match
Best Local Similarity 2.9%; Score 101.5; DB 7; Length 148;
Matches 37; Conservative 18; Mismatches 57; Indels 19; Gaps 5;

QY 491 TDSQKVNRTLDLDTVILEADGM-----SYMDALYAN--KASLVAIFNIIDAD 541
DB 5 TEBQIAEFKAPSLFDK-----GDGTTTKELGTVMSLQNPTEARLQDINVEDAD 58

QY 542 NSGEITLDEFEFTAIDLVAHMPGAVSKAMLEKCRMDLNGDKVDLNEFLEAF-RLSDDL 600
DB 59 GNGTTIDPEPFLT---MMARKKQDTSSEBIRAFVFDKDGNGYISAELRHVMTNLGK 115

QY 601 HRKEQDENIR 611
```


Query Match 2.8%; Score 98; DB 6; Length 5335;

Best Local Similarity 18.9%; Pred. No. 1.2e+02;

Matches 154; Conservative 103; Mismatches 265; Indels 292; Gaps 43;

9 AIFQKVRHQRREMR-----CWMQIF--ONLEVASQ-- 43

4364 ALLEQK--HVSSKMEERKSKLEBALNATEFQNSLOEFINWTLAEOSLNIASPSSL 4420

44 -----DOALYKPF--NDLIKMPQ-----AAGRNOYOGSAH-----VSU----- 77

4421 ILNTVLSQIEBHKVFAFENVAHNRDQIIEIDQNGQLKFLSQKQDVILKNLVSVOGRWE 4480

78 -----LDDKODLVEERFGDIYNAKIE--LPIRKNHIDLIDV----- 111

4481 KVVQRSIERGRSLDDAKRAKQFHEAMKCLIDWLEDBESHLDBELISNDPDKIKLOLSK 4540

112 ---PRKRGRLHPKYVALILREAKSLKQ---LPNISPVSTAVSQV---TVCG--- 157

4541 HKEFQKTLGGK--QPVYDTTI--RTGRALKEKTLIPEDTQKLDNPLFGEVRDKMDTVCGKSV 4597

158 DLHGKLDLVLVHKNGLPSSSNPYVNGDFVDRGR-----GLB 197

4598 BROHKEBAL-----LPSGQPMALQALVDMLYKVEPOLAEDQPVHGDL 4642

198 VLLILSLYLAFPNV-----FLNRGNHDSVMNR-----YGFIREV----- 235

4643 LVNMLDAHVKVFOKELGKRTGVQLKRSGR--LIENSRODTTWKQLOBELSTRMDTVCG 4701

236 -----ESKYPNRNKRILAFIDVY---RWLPLGSVANSRLVHGSPDSTSLDKSI 286

4702 KLSVSKOSRLBQALKQAEVFRDVTVMLEWLSA---BQTLRFGALPDPTA--LQSLI 4756

287 DRGKYVSLRPPLDGPRLDTEMOQIFDIWMSDPOATMG-----CVPTLRGAGVWFG 340

4757 D-----THKEFMKVEBKRV--DV---NSAAMGEVILAVCHPDCITTIKMI- 4799

341 PDVTNPLQHRISLVYRSHCEKPNHGFMDNKIITIFASNYAIGSNKAGAYIRLNQ 400

4800 -----TIRARFEVUTMAK-----OHQRLLETALS------LVANAE 4832

401 LMPHFVQYISAAGT-----KRLSPQRMGIVESSALKELAVMRDHRD 444

4833 LLEBLAMIQWAEFTLLIQRODEPIPNIDRVKAL-----IAHQTFMEEMTKQDQVD 4885

445 ELDEBPRK--YDPKDSGYI--SISHMCKMENVTKGLP----- 479

4886 RVYTKYRKNIIEPTHAPEIKSRSGRKSLSQTPPPMPTLISSEAKNPRINQISARMOQ 4945

480 -WRLLRDKLAPGTSOKVNNRTLDLDTVILEABD-----GMSVMDALYANKASLV 532

4946 VMLLALER-----QRKLN-----DALDRLEELKEFANFPDVRKKYKMMNHNKSRVM 4994

533 AIFNIIDADNSGRTTIDDEFETAIID-LVVAHMPGAYSAEMLKRCRMDLNGDKVDLNEF 591

4995 DFFRRIDKDDGKITRQEF---IDGILASKRP--TTKLEMTAVADIPDRDGDYIDYEF 5049

592 LEAFRLS--DLHRKEQDENIRRRSTGRPSVAKTA 624

5050 VAAHPKDAVRPTTDADKIEDVTRQVACCKA 5083

RESULT 31

US-10-995-561-774

; Sequence 774; Application US/10995561

; Publication No. US20050272054A1

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

; FILE REFERENCE: C001559

; CURRENT APPLICATION NUMBER: US/10/995,561

CURRENT FILING DATE: 2004-11-24

; NUMBER OF SEQ ID NOS: 85702

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 774

; LENGTH: 5406

; TYPE: PRN

; ORGANISM: Homo sapiens

US-10-995-561-774

Query Match 2.8%; Score 98; DB 6; Length 5406;

Best Local Similarity 18.9%; Pred. No. 1.2e+02;

Matches 154; Conservative 103; Mismatches 265; Indels 292; Gaps 43;

9 AIFQKVRHQRREMR-----CWMQIF--ONLEVASQ-- 43

4435 ALLEQK--HVSSKMEERKSKLEBALNATEFQNSLOEFINWTLAEOSLNIASPSSL 4491

44 -----DOALYKPF--NDLIKMPQ-----AAGRNOYOGSAH-----VSU----- 77

4492 ILNTVLSQIEBHKVFAFENVAHNRDQIIEIDQNGQLKFLSQKQDVILKNLVSVOGRWE 4551

78 -----LDDKODLVEERFGDIYNAKIE--LPIRKNHIDLIDV----- 111

4552 KVVQRSIERGRSLDDAKRAKQFHEAMKCLIDWLEDBESHLDBELISNDPDKIKLOLSK 4611

4612 HKEFQKTLGGK--QPVYDTTI--RTGRALKEKTLIPEDTQKLDNPLFGEVRDKMDTVCGKSV 4668

112 ---PRKRGRLHPKYVALILREAKSLKQ---LPNISPVSTAVSQV---TVCG--- 157

4612 HKEFQKTLGGK--QPVYDTTI--RTGRALKEKTLIPEDTQKLDNPLFGEVRDKMDTVCGKSV 4668

158 DLHGKLDLVLVHKNGLPSSSNPYVNGDFVDRGR-----GLB 197

4669 BROHKEBAL-----LPSGQPMALQALVDMLYKVEPOLAEDQPVHGDL 4713

198 VLLILSLYLAFPNV-----FLNRGNHDSVMNR-----YGFIREV----- 235

4714 LVNMLDAHVKVFOKELGKRTGVQLKRSGR--LIENSRODTTWKQLOBELSTRMDTVCG 4772

236 -----ESKYPNRNKRILAFIDVY---RWLPLGSVANSRLVHGSPDSTSLDKSI 286

4773 KLSVSKOSRLBQALKQAEVFRDVTVMLEWLSA---BQTLRFGALPDPTA--LQSLI 4827

287 DRGKYVSLRPPLDGPRLDTEMOQIFDIWMSDPOATMG-----CVPTLRGAGVWFG 340

4828 D-----THKEFMKVEBKRV--DV---NSAAMGEVILAVCHPDCITTIKMI- 4870

341 PDVTNPLQHRISLVYRSHCEKPNHGFMDNKIITIFASNYAIGSNKAGAYIRLNQ 400

4871 -----TIRARFEVUTMAK-----OHQRLLETALS------LVANAE 4903

401 LMPHFVQYISAAGT-----KRLSPQRMGIVESSALKELAVMRDHRD 444

4904 LLEBLAMIQWAEFTLLIQRODEPIPNIDRVKAL-----IAHQTFMEEMTKQDQVD 4956

445 ELDEBPRK--YDPKDSGYI--SISHMCKMENVTKGLP----- 479

4957 RVYTKYRKNIIEPTHAPEIKSRSGRKSLSQTPPPMPTLISSEAKNPRINQISARMOQ 5016

480 -WRLLRDKLAPGTSOKVNNRTLDLDTVILEABD-----GMSVMDALYANKASLV 532

5017 VMLLALER-----QRKLN-----DALDRLEELKEFANFPDVRKKYKMMNHNKSRVM 5065

533 AIFNIIDADNSGRTTIDDEFETAIID-LVVAHMPGAYSAEMLKRCRMDLNGDKVDLNEF 591

5066 DFFRRIDKDDGKITRQEF---IDGILASKRP--TTKLEMTAVADIPDRDGDYIDYEF 5120

592 LEAFRLS--DLHRKEQDENIRRRSTGRPSVAKTA 624

5121 VAAHPKDAVRPTTDADKIEDVTRQVACCKA 5154

RESULT 32

US-10-995-561-779

; Sequence 779; Application US/10995561


```

; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 779
; LENGTH: 5415
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-995-561-779

Query Match      2.8%; Score 98; DB 6; Length 5415;
Best Local Similarity 18.9%; Pred. No. 1.2e+02;
Matches 154; Conservative 103; Mismatches 265; Indels 292; Gaps 43;

QY      9 AITIOKTYRRHQAARRMQR-----CNWQIF--QNTLYASBQ-- 43
DB      4456 ALLEQKV---HYVSSKMERKSKLEBALNLATEFQNSLOEFINWLTLAGSLINIASPPSL 4512

QY      44 -----DQAELYKPF--NDLIKMPQ-----AAGRKNOYGSAH-----VSU----- 77
DB      4513 ILMTVLSQIEBKHFVAFNVAHNDQIIEIDQTGNQKFLSOKODVVLIRNLLVSVQSRWE 4572

QY      78 -----LDDKODLVEBFGDLYVNAKIE--LPIRKNHIDLLIDV----- 111
DB      4573 KVVQRSIERGRSLDDAKRAKQFHEAMKCLIDWLEDSHLSBELISNDPKIKLOLSK 4632

QY      112 ---FRKRGRLHPKYVALILREAAKSLQ---LPIISPVSTAVSQV-----TVCG--- 157
DB      4633 HKEFQKTLGSK-QPYVDITI--RTGRALKETKLLPEDTQKLNFLGEVNDKMDTVGSKV 4689

QY      158 DLHGKIDLLVLVHKNGLPSSSNPYFNGDFVDRGR-----GLE 197
DB      4690 ERQHKLEAL-----LFSGQFMDALQALVDMLYKVEPOLAEDQPVHGDL 4734

QY      198 VLLLLSLYLAFPNV-----FLNRGNHDSVNNAR-----YGFIREV----- 235
DB      4735 LVNMLDAHKVFOKEIGKRTGVQLKRGRE--LIENSRDQTTWVGQLOELSTRMDTVG 4793

QY      236 -----ESKTPRNKRIILAFIDEVY---RWLPLGSLVNSRVLLVHGSPDSTSLDIKSI 286
DB      4794 KLSVSKQSLQALKAELFRDVTVMLEWLSA---EQTLRFRALPDDEA--LQSLI 4848

QY      287 DRGKYVSLRPLTDEBPLDKTEWQOIFDIMMSDPQATWG-----CVPTLIRGAGVWFG 340
DB      4849 D-----THKEFMKKVEKRV-DV---NSAVAMGEVILAVCHPDCITTIKHWI- 4891

QY      341 PDVTNDFLOHRRLSYVIRSHCEKPNGHEFMNDKIIITISASNYVALIGSNKAYIYRLNQ 400
DB      4892 -----TIIRAREEVLTWAK-----OHQORLETALE-----LVANAE 4924

QY      401 LMPHFVOYISAASQT-----KRLSFQRMGIYESSALKELAVRMDHRD 444
DB      4925 LLEBLAMIQMAETTLIQDOEPIDPQNIIDRYKAL-----IAEHQTFMEBMTKQPDVD 4977

QY      445 ELDEBFRK--YDPKDSGYI--SISHMCKYMENTYKGLP----- 479
DB      4978 RVTKTYKRNIEPTHAPFIEKSRGKSLSQPTPPMPLISQSEAKNPRIQLSARMOQ 5037

QY      480 -WRLRLDKLAGTDSQVNTYRITDLDIVILAEABD-----GMSVMDALYANKASLV 532
DB      5038 VMLALAEK-----QRGLN-----DALDRLEELKEFANPDFVWRKCKYMMAMHAKSRVA 5086

QY      533 AIFNIIDADNSGKITLDEFTALID-LVAHMPGAYSAEMLEKCRMDLNGDGKVDLNEF 591
DB      5087 DFFRIDKDDGKITRQEF--IDGILASKFP--TTLKMTAVADIDRDDGDYIDYEF 5141

```

```

QY      592 LBAFRLS-DLHRKEQODENIRRRSGTGPVAXTA 624
DB      5142 VAAHFNKDAVPTTDADKIEBVTROVAQCKCA 5175

RESULT 33
US-10-995-561-775
; Sequence 775, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 775
; LENGTH: 5464
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-995-561-775

Query Match      2.8%; Score 98; DB 6; Length 5464;
Best Local Similarity 18.9%; Pred. No. 1.2e+02;
Matches 154; Conservative 103; Mismatches 265; Indels 292; Gaps 43;

QY      9 AITIOKTYRRHQAARRMQR-----CNWQIF--QNTLYASBQ-- 43
DB      4456 ALLEQKV---HYVSSKMERKSKLEBALNLATEFQNSLOEFINWLTLAGSLINIASPPSL 4512

QY      44 -----DQAELYKPF--NDLIKMPQ-----AAGRKNOYGSAH-----VSU----- 77
DB      4513 ILMTVLSQIEBKHFVAFNVAHNDQIIEIDQTGNQKFLSOKODVVLIRNLLVSVQSRWE 4572

QY      78 -----LDDKODLVEBFGDLYVNAKIE--LPIRKNHIDLLIDV----- 111
DB      4573 KVVQRSIERGRSLDDAKRAKQFHEAMKCLIDWLEDSHLSBELISNDPKIKLOLSK 4632

QY      112 ---FRKRGRLHPKYVALILREAAKSLQ---LPIISPVSTAVSQV-----TVCG--- 157
DB      4633 HKEFQKTLGSK-QPYVDITI--RTGRALKETKLLPEDTQKLNFLGEVNDKMDTVGSKV 4689

QY      158 DLHGKIDLLVLVHKNGLPSSSNPYFNGDFVDRGR-----GLE 197
DB      4690 ERQHKLEAL-----LFSGQFMDALQALVDMLYKVEPOLAEDQPVHGDL 4734

QY      198 VLLLLSLYLAFPNV-----FLNRGNHDSVNNAR-----YGFIREV----- 235
DB      4735 LVNMLDAHKVFOKEIGKRTGVQLKRGRE--LIENSRDQTTWVGQLOELSTRMDTVG 4793

QY      236 -----ESKTPRNKRIILAFIDEVY---RWLPLGSLVNSRVLLVHGSPDSTSLDIKSI 286
DB      4794 KLSVSKQSLQALKAELFRDVTVMLEWLSA---EQTLRFRALPDDEA--LQSLI 4848

QY      287 DRGKYVSLRPLTDEBPLDKTEWQOIFDIMMSDPQATWG-----CVPTLIRGAGVWFG 340
DB      4849 D-----THKEFMKKVEKRV-DV---NSAVAMGEVILAVCHPDCITTIKHWI- 4891

QY      341 PDVTNDFLOHRRLSYVIRSHCEKPNGHEFMNDKIIITISASNYVALIGSNKAYIYRLNQ 400
DB      4892 -----TIIRAREEVLTWAK-----OHQORLETALE-----LVANAE 4924

QY      401 LMPHFVOYISAASQT-----KRLSFQRMGIYESSALKELAVRMDHRD 444
DB      4925 LLEBLAMIQMAETTLIQDOEPIDPQNIIDRYKAL-----IAEHQTFMEBMTKQPDVD 4977

QY      445 ELDEBFRK--YDPKDSGYI--SISHMCKYMENTYKGLP----- 479
DB      4978 RVTKTYKRNIEPTHAPFIEKSRGKSLSQPTPPMPLISQSEAKNPRIQLSARMOQ 5037

```


ORGANISM: Mus musculus
US-10-118-590-6

Query Match
Best Local Similarity 20.8%; Score 97; DB 6; Length 216;
Matches 50; Conservative 33; Mismatches 69; Indels 88; Gaps 9;

424 MGIVSSALKEKLAARMRHRDELDEPRKYPKOSGIYSISHCKWMENTYKGLPRL 483
5 MGRSSSLQTKQ-----RRSPKIDBE-----LEMTVCHREPGLEQLAQNFT 49
484 RDKLAPGDSQKVVNRTLDLDTVILEADGWSVDALYA-----NKASLYA--IFN 536
50 KREL-----QVLYRGKACPCSGVNE-----ETFKQIYAQPPPGDASTYAHYFN 96
537 IIDANGSEITLDEFEIADLLV-----AMPGAYSXAEMLKCR----- 576
97 AFDTTQTSVAFEDVFTALSTLKGTVHKLRTFNLVDINKDGYINKEMWDIYKAIYD 156
577 -----MDLNGDKYDINEFLAFLSLDLHREQDENIR 612
157 MNGKTYPVAKEDTPRQHDVDFQKMDKNKGIVTLDEFLSC-----QBDNIMR 207

RESULT 39
US-10-878-556A-161
Sequence 161, Application US/10878556A
Publication No. US20050266399a1
GENERAL INFORMATION:
APPLICANT: Hoffmann La-Roche Inc.
TITLE OF INVENTION: HCV regulated protein expression
FILE REFERENCE: 21762
CURRENT APPLICATION NUMBER: US/10/878, 556A
CURRENT FILING DATE: 2004-06-28
NUMBER OF SEQ ID NOS: 199
SOFTWARE: PatentIn version 3.1
SEQ ID NO 161
LENGTH: 814
TYPE: PRT
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: sw_hum/f39_human
DATABASE ENTRY DATE: 1997-11-01
US-10-878-556A-161

Query Match
Best Local Similarity 19.8%; Score 97; DB 6; Length 814;
Matches 73; Conservative 53; Mismatches 119; Indels 124; Gaps 16;

185 NGDF---VDRGKGLLEVLILLLSLYLAFPAVFLNRGNHEDSVNNAHYGFIRVESKYP 240
522 NGDYLCVVDRTPKGTQGVVT-----NPEIFNMRKQVVDVEMK-- 562
241 RNHRILAFIDBYRMLPLGSVLNSRVLIYNG-----GFSDSLSLLIKSIDRGK 290
563 ---FTILAF-----AMEPNG---SKFVLHGEAPRISVSFYHYKNGKIELIKMPX-- 608
291 YVSLIRPLTDEBPLDKTEMQOIFDIMWSDPOATMGCVPTNLRAAGVFGPDVTNFIQR 350
609 -----QANTIFWS-POQGFVTLAAGLRSMNGALAFVDTSDCYMN 647
351 HRLSVYIRSHCKENGHEFMEDNKITITFS-----ASNYTVAIGSNKAYIRLANQMPHF 405
648 IAEHYMASDVEWDPFG-----RYVYTSVSWMSHKVDNAVYMLTFOGRLQKNK--DRF 699
406 VQY-----ISASQTKRLS-----FKORMGIVESSALKELAVRRDRHDELD 448
700 COLMRPPPTLLSQEQIKQIKDLKYSKIFEQDRLSQSKASKELEVERR---TMBE 755
449 EFRKIDPKDSGYISISHCKWMENTYKGLPRLRLDKLAPGDSQKVVN-----RTL 502
756 DFRK-----RKMAQELYMEOGKNERLRLRGSVDTDLSDSVDDMBERTI 799

Query Match
Best Local Similarity 20.8%; Score 97; DB 7; Length 1184;
Matches 157; Conservative 88; Mismatches 257; Indels 252; Gaps 41;

503 DLDTDVIL 511
800 EFPVTEBIT 808

RESULT 40
US-11-115-639-50
Sequence 50, Application US/11115639
Publication No. US20050282242A1
GENERAL INFORMATION:
APPLICANT: Roche, David
APPLICANT: Murphy, Christopher
APPLICANT: MacNeil, Ian
TITLE OF INVENTION: SCREENING ASSAYS FOR ANTIMICROBIAL
TITLE OF INVENTION: AGENTS
FILE REFERENCE: 50150/075003
CURRENT APPLICATION NUMBER: US/11/115, 639
CURRENT FILING DATE: 2005-04-27
PRIOR APPLICATION NUMBER: 60/566, 858
PRIOR FILING DATE: 2004-04-30
PRIOR APPLICATION NUMBER: 60/565, 679
PRIOR FILING DATE: 2004-04-27
NUMBER OF SEQ ID NOS: 55
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 50
LENGTH: 1184
TYPE: PRT
ORGANISM: L. monocytogenes
US-11-115-639-50

Query Match
Best Local Similarity 20.8%; Score 97; DB 7; Length 1184;
Matches 157; Conservative 88; Mismatches 257; Indels 252; Gaps 41;

16 YRRHQAEREMQRCNMQIFONTLEYAS--EODQAEIYFNDLTKHMPQAARKKQYOGSA 73
11 YGRHRTKRSFAR-----ISEVLELPNILEIQTASVCMFDEGLAREMRDSDPIEDFAGNL 65
74 HVSVLDDKDDLVBEFGDIYNAKI ELPTRKNIHDLIDVFRKKRGNRLHPRYVALILREAA 133
66 SLRPID-----YDLGEPKYSVEESKN-----RANVAAPLRVGLRL----- 101
134 KSLKQLPNISPVSTAYSQVTVCGD--LHKLDLVLVYHKG-----LPSSSPHY 183
102 -----INKETGEVVDQEVFMGDPPLMTMGTFIT--NGAERYVSQLVRSPEGV- 148
184 FNGDPVDRGKRGLEVLILLLSLYLAFPAVFLNRG-----NHEDSVNNAH----- 228
149 FNGKLDNRNGKKG-----FQSTVIYPRNGAMLEYETDAKDVAHVIRIDRTKLPV 195
229 -----YGFIREV-----ESKYPRN-----HKRIILAFIDBYRMLPLG--SV 262
196 TVLLRALGFGSDQDEIIDLDGNDYLRNTLEKMDTNAEKALL-----EYRLRARGBPPTV 251
263 LNSRVLIYHGGFSDSTSLDIKSIDRGK-----YVSLIRPLT----- 301
252 DNARSLLV--SRFPDPKRYDL--ASVGRYKINKKHLKNRLLFNQTLAETLVLPETGEIYASK 309
302 GBPLDKTEMQOIFDIMWSDPOATMGCVPTNLRAAGV--WFGPDVTNFIQHRHSYVIRSH 360
310 GDILDRNLDOI-----IPNLENGVGFRTLRP--TDGMEBSVLVQSIKITIY 353
361 ECKPENGHEFMEDNKITITFSASNYTVAIGSNKAYIRLANQMPHFVQYISASQTKRLS- 419
354 --APNDEB-----KEINI-----IGN--AYIEEN-----VGHITPSDIISSISY 368
420 -FKORMGIVESSALKELA--VRMDHDELDEPRKYPKOSGIYSISHCKW-----ME 471
389 FFINLHGVGDTDOI DILGNRLRSVGBLLQNP-----IGLSRMRVVRBERMSIQ 439
472 NVTKLGAPRL-----LRDKLAPGDSQKVVNRTLDLDTVILEA-----E 514
440 DMTTI--TPQOLINIRPVVASIKERFGSSQLSQFMNQNPDLDELTYKRLSALRGGLTRE 498

```

Qy 515 ADGMSVMDALYANKASU-----VAIFNIIDADNSGEITLDER-----E 552
Db 499 RAGYEVARDVHYSHYGRMCPIETPEGNIGLFTYL---SSPAKVNRKGPIETPYRRVDPB 554
Qy 553 T-----AIDLVAHMPGAYSKAMLEKCRMDLNGDKVDLNEFLAFLSLDIHRKQOD 607
Db 555 TNRVTDKIDYLTADDEEDNVVVAQANSK-----LDEQOTFTRBEVMAFRSENL---AVEK 606
Qy 608 ENIRRRSTGRPSVAKTATDPVTLLADKISNTLY 641
Db 607 ERIDYMDVSPKQVSVATACIPLENDSDSRALM 640

```

Search completed: January 20, 2006, 20:04:22
 Job time : 37 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2006, 19:39:22 ; Search time 40 Seconds
(without alignments)
1589.982 Million cell updates/sec

Title: US-09-463-733-1

Perfect score: 3442

Sequence: 1 MDENAIRAIFQKWTRRHQ.....VEHDDIDPTDCESKVIDPKKS 661

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: _pir1:*
2: _pir2:*
3: _pir3:*
4: _pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 3442 | 100.0 | 661 | 1 A42287 | phosphoprotein pho |
| 2 | 1317 | 38.3 | 707 | 2 T42239 | probable phosphor |
| 3 | 1305 | 37.9 | 722 | 2 T34072 | hypothetical prote |
| 4 | 593.5 | 17.2 | 704 | 2 T14614 | hypothetical prote |
| 5 | 569 | 16.5 | 499 | 1 A55346 | phosphoprotein pho |
| 6 | 567.5 | 16.5 | 498 | 1 S52570 | phosphoprotein pho |
| 7 | 558 | 16.2 | 526 | 1 T45058 | phosphoprotein pho |
| 8 | 557.5 | 16.2 | 533 | 2 B48058 | phosphoprotein pho |
| 9 | 541 | 15.7 | 479 | 1 T46076 | phosphoprotein pho |
| 10 | 519 | 15.1 | 473 | 1 T40391 | phosphoprotein pho |
| 11 | 479.5 | 13.9 | 513 | 1 S52571 | phosphoprotein pho |
| 12 | 444 | 12.9 | 413 | 2 T51811 | phosphoprotein pho |
| 13 | 441 | 12.8 | 326 | 2 T09995 | phosphoprotein pho |
| 14 | 436.5 | 12.7 | 314 | 2 S52371 | phosphoprotein pho |
| 15 | 435.5 | 12.7 | 327 | 2 A32550 | phosphoprotein pho |
| 16 | 434 | 12.6 | 729 | 2 T09913 | probable phosphor |
| 17 | 433.5 | 12.6 | 326 | 2 T09544 | phosphoprotein pho |
| 18 | 429.5 | 12.5 | 323 | 2 A32549 | phosphoprotein pho |
| 19 | 429 | 12.5 | 312 | 2 B32550 | phosphoprotein pho |
| 20 | 428 | 12.4 | 328 | 2 S20882 | phosphoprotein pho |
| 21 | 426.5 | 12.4 | 317 | 2 T03594 | phosphoprotein pho |
| 22 | 425 | 12.3 | 401 | 1 T21288 | phosphoprotein pho |
| 23 | 424.5 | 12.3 | 312 | 2 S31086 | phosphoprotein pho |
| 24 | 421 | 12.2 | 323 | 2 T09550 | phosphoprotein pho |
| 25 | 421 | 12.2 | 326 | 2 T09548 | phosphoprotein pho |
| 26 | 420.5 | 12.2 | 317 | 2 T03304 | probable phosphor |
| 27 | 420 | 12.2 | 291 | 1 T21322 | phosphoprotein pho |
| 28 | 419.5 | 12.2 | 321 | 2 S31088 | phosphoprotein pho |
| 29 | 418.5 | 12.2 | 310 | 1 T25593 | phosphoprotein pho |

| | | | | | |
|----|-------|------|-----|----------|--------------------|
| 30 | 418.5 | 12.2 | 312 | 2 S32595 | phosphoprotein pho |
| 31 | 418.5 | 12.2 | 312 | 2 S24264 | phosphoprotein pho |
| 32 | 418.5 | 12.2 | 325 | 2 T09547 | phosphoprotein pho |
| 33 | 413.5 | 12.0 | 401 | 2 T18936 | phosphoprotein pho |
| 34 | 413 | 12.0 | 302 | 1 PAFPIA | phosphoprotein pho |
| 35 | 412 | 12.0 | 316 | 2 S26225 | phosphoprotein pho |
| 36 | 411.5 | 12.0 | 308 | 1 PABY3 | phosphoprotein pho |
| 37 | 411.5 | 12.0 | 333 | 1 T16476 | phosphoprotein pho |
| 38 | 409.5 | 11.9 | 382 | 1 T19701 | phosphoprotein pho |
| 39 | 409 | 11.9 | 316 | 2 S29317 | phosphoprotein pho |
| 40 | 408 | 11.9 | 312 | 2 S31089 | phosphoprotein pho |
| 41 | 408 | 11.9 | 322 | 1 S31087 | phosphoprotein pho |
| 42 | 407.5 | 11.8 | 323 | 2 T76572 | phosphoprotein pho |
| 43 | 407.5 | 11.8 | 337 | 1 C32550 | phosphoprotein pho |
| 44 | 406.5 | 11.8 | 337 | 1 T76573 | phosphoprotein pho |
| 45 | 405.5 | 11.8 | 323 | 1 S35699 | phosphoprotein pho |

ALIGNMENTS

RESULT 1
A42287
phosphoprotein phosphatase (BC 3.1.3.16) rdgc - fruit fly (Drosophila melanogaster)
N:Alternate names: retinal degeneration protein C
C:Species: Drosophila melanogaster
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Jul-2004
C:Accession: A42287; S27811
R:Steele, F. R.; Washburn, T.; Rieger, R.; O'Touse, J. B.
Cell 69, 669-676, 1992
A:Title: Drosophila retinal degeneration C (rdgc) encodes a novel serine/threonine prot
A:Reference number: A42287; MUID:9226398; PMID:1316807
A:Accession: A42287
A:Molecule type: DNA
A:Residues: 1-661 <ST>
A:Cross-references: UNIPROT:P40421, UNIPARC:UPI000000728A; EMBL:W69628; NID:G158237; PFI
A>Note: sequence extracted from NCBI backbone (NCBIN:103017, NCBIN:103023, NCBIN:103028,
, NCBIN:103104, NCBIPI:103001)
C:Genetics:
A:Gene: rdgc
A:Cross-references: FlyBase:FBgn0004366
A:Introns: 5/1; 47/3; 74/3; 116/3; 172/3; 240/3; 290/3; 312/3; 374/3; 510/3; 544/1
C:Function:
A:Description: catalyzes the hydrolytic dephosphorylation of protein-phosphoserine and p
A>Note: prevents light-induced retinal degeneration
C:Superfamily: serine/threonine protein phosphatase with BF-hands; calmodulin repeat hom
C:Keywords: calcium binding; BF hand; iron; metalloprotein; nucleus; phosphoric monoeste
F:120-401/Domains: phosphoprotein phosphatase homology <PP>
F:152-221/Domains: phosphoesterase core homology <EP>
F:1526-558/Domains: calmodulin repeat homology <EP>
F:156-598/Domains: calmodulin repeat homology <EP>
F:158-160,187/Binding site: iron (Aep, His, Asp) #status predicted
F:187-219,271,360/Binding site: zinc (Asp, Asn, His, His) #status predicted
F:190,220,384/Active site: Asp, His, Tyr #status predicted
F:191,333/Binding site: substrate phosphate (Arg) #status predicted

Query Match 100.0%; Score 3442; DB 1; Length 661;
Best Local Similarity 100.0%; Pred. No. 1.5e-208;
Matches 661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 1 | MDENAIRAIFQKWTRRHQARRRQRCNMOIFONTLBYASQDAELYKFPNDLIKMP | 60 |
| DB | 1 | MDENAIRAIFQKWTRRHQARRRQRCNMOIFONTLBYASQDAELYKFPNDLIKMP | 60 |
| QY | 61 | QAAGKKNQYQSSAHVSDDDKODLVEEFGDIYNAKIELEPIRKNIHDLIDVRKKRGNTL | 120 |
| DB | 61 | QAAGKKNQYQSSAHVSDDDKODLVEEFGDIYNAKIELEPIRKNIHDLIDVRKKRGNTL | 120 |
| QY | 121 | HKYVALILIRBAKSLKQLEPNISPVSTAVSQQVTCGLHGLDLDLVLLHKGKLPSSSN | 180 |
| DB | 121 | HKYVALILIRBAKSLKQLEPNISPVSTAVSQQVTCGLHGLDLDLVLLHKGKLPSSSN | 180 |
| QY | 181 | PYVENGDFVRDGRGLVLLILSLYLAEPNAVPLNRGNHEDSVNNAKRGPIREVSXTP | 240 |

```
Db 181 PVTFFNGDFPRDKRGKGVLLLLSLYLAFPAVFLNRGNHDSVMNAKGFIREVESKTP 240
Qy 241 RNHKKILAFIDEVYRMLPLGSLVNSRVLIYHGGFSDSTSLDLKSIDRGKVSILRPPLT 300
Db 241 RNHKKILAFIDEVYRMLPLGSLVNSRVLIYHGGFSDSTSLDLKSIDRGKVSILRPPLT 300
Qy 301 DDEPLDKTEWQOIFDIWMSDPOATMGCVPTLRGAGVWGPRVTNPLQRHLSVYISHS 360
Db 301 DDEPLDKTEWQOIFDIWMSDPOATMGCVPTLRGAGVWGPRVTNPLQRHLSVYISHS 360
Qy 361 ECKPNGHEFMHDKITITISASNYAIGSNKGAAYIRLNQMLPHFYQYISAASQTKRLSF 420
Db 361 ECKPNGHEFMHDKITITISASNYAIGSNKGAAYIRLNQMLPHFYQYISAASQTKRLSF 420
Qy 421 KQRMGIVBSSALKEIAYVRMRDRDELDEFKRYDPPDSGYISISHMKVMENVTGLGJPM 480
Db 421 KQRMGIVBSSALKEIAYVRMRDRDELDEFKRYDPPDSGYISISHMKVMENVTGLGJPM 480
Qy 481 RLRLDKLAPGTSOKVNVARTLDDLDTQVYILAEADGMSVMALYANKASLVAIFNIIDA 540
Db 481 RLRLDKLAPGTSOKVNVARTLDDLDTQVYILAEADGMSVMALYANKASLVAIFNIIDA 540
Qy 541 DNSGRIITLDEFETALDLVAHMPGAYSKAMLEKCMMDLNGDKVDLNEFLAERLSD 600
Db 541 DNSGRIITLDEFETALDLVAHMPGAYSKAMLEKCMMDLNGDKVDLNEFLAERLSD 600
Qy 601 HKKQODENIRRRSTGRPSVAATADPTVLADKISKNTLVVEHDIPTDCSKVIDPPK 660
Db 601 HKKQODENIRRRSTGRPSVAATADPTVLADKISKNTLVVEHDIPTDCSKVIDPPK 660
Qy 661 S 661
Db 661 S 661
```

RESULT 2

142239
Probable phosphoprotein phosphatase (BC 3.1.3.16) - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 12-Jul-2004
C/Accession: T42239
R:Sherman, P. M.; Sun, H.; Macle, J. P.; Williams, J.; Smallwood, P. M.; Nathans, J.
Proc. Natl. Acad. Sci. U.S.A. 94, 11639-11644, 1997
A/Title: Identification and characterization of a conserved family of protein serine/thr
A/Reference number: Z22115; MUID:97471020; PMID:9326663
A/Accession: T42239
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-707 <SHE>
A/Cross-references: UNIPROT:O01921; UNIPARC:UP10000060867; EMBL:AF023454; NID:92586408;
C/Superfamily: serine/threonine protein phosphatase with EF-hand; calmodulin repeat hom
C/Keywords: EF hand; phosphoric monoester hydrolase; serine/threonine-specific phosphatase

```
Query Match 38.3%; Score 1317; DB 2; Length 707;
Best Local Similarity 42.1%; Pred. No. 5.3e-75;
Matches 266; Conservative 116; Mismatches 200; Indels 50; Gaps 12;

Qy 6 IBAALFIQMYRRHQARRMQRCKNQIFONTLEYASEDOAEIYKFENDLIKMPQOAGR 65
Db 83 IKSALILQKMYRRCERLARARRATQIFALAEAGEOQLKRYVITLYDFADVIR 142
Qy 66 KNOYQG--SAHVSVLDD--KDLVEEFGDIYNAKIF-----LPIRK 102
Db 143 GGVENGRNSPLMSALSHYAKPSLMSRGFTVKMLLEDTPTNVDIDRNYKGPFLSLPDK 202
Qy 103 NHIDILIDVFRKKRGRLAPKYVALILREASLSKQLPMSIPVSTAVSQOVTAVCGDLNKK 162
Db 203 PQVAKMTEAF--KVNKVLHPKYVLMILHEARKIFKAMPVSRSRISTISNQVTICGDLNKK 260
Qy 163 LDPLVLVLRKNGILPSSSNPYVFNQDFVDRGKGLVYLILLSLYLAFPAVFLNRGNHSD 222
Db 261 PDDLCLITLVKNGIPSVNPIYIFNQGDFVDRGGGSIIVLCVLPALVIVDPMSITLNRGNHSD 320
```

```
Qy 223 SVMNRRYGFIREVESKTPRNHKKILAFIDEVYRMLPLGSLVNSRVLIYHGGFSDSTSLDL 282
Db 321 HINMLRYGFIREVESKTPRNHKKILAFIDEVYRMLPLGSLVNSRVLIYHGGFSDSTSLDL 380
Qy 283 IKSIDRGKVSILRPPLTDLGEBLTK-----TEMQOIFDIWMSDPOATMGCVPTLRGA 335
Db 381 LDKTFRHRFQSVLPVPVAKGSESEKNSAVAVDEKQMLDIWMSDPOATMGCVPTLRGA 440
Qy 336 GVGFGPDVTDNPLQRHLSVYIRSHCKPNGBFMHDKITITISASNYAIGSNKGAAYI 395
Db 441 GSYGADITAFLEHGGRLLVVRHGECKFEYEFSSHNTCLTVFSASNYIETGSRGAYV 500
Qy 396 R-LNQLMPPHYQYISAASQTKRLS-FQRMGIYBSSALKEIAYVRMRDRDELDEFKRY 453
Db 501 KFIGSKQPHFYQYM--ASKTKRSTLNERGVVBSAVKELKETSFPHTDLOKFEIIM 558
Qy 454 DPKQSGYISISHMKVMENVTGLGJPMRLTLDKLP--GTDGOKVNVARTLDDLDTQVILE 512
Db 559 DIKSGKPLIKMSDQVRRITGLNPLIALPRTATLSBDGKIVMYKEDRAIAYGGTHA 618
Qy 513 AEADGMSVMALYANKASLVAIFNIIDANSGETLDEFETALDLVAHMPGAYSKAM- 571
Db 619 QEKD---IVESLYHRKSTLTLTFRPMQDNQSGVSMKEFIDACEVL-----GKVTKRPLQ 670
Qy 572 ----LEKCMMDLNGDKVDLNEFLAERLSD 599
Db 671 TDYISQIAESIDFNKDPFDILNELLEAFRLVD 702
```

RESULT 3

734072
Hypothetical protein F23H11.8 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 12-Jul-2004
C/Accession: T34072
R:Roehling, T.; Wohldmann, P.
Submitted to the EMBL Data Library, May 1997
A/Description: The sequence of C. elegans cosmid F23H11.
A/Reference number: Z21470
A/Accession: T34072
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-722 <ROH>
A/Cross-references: UNIPROT:O01921; UNIPARC:UP10000175983; EMBL:AF003389; PIDN:AACT1139.
A/Experimental source: strain Bristol N2; clone F23H11
C/Genetics:
A/Gene: CESP:F23H11.8
A/Map position: 3
A/Intons: 91/1; 139/3; 353/3; 579/1; 674/3
C/Superfamily: serine/threonine protein phosphatase with EF-hand; calmodulin repeat hom
C/Keywords: EF hand

```
Query Match 37.9%; Score 1305; DB 2; Length 722;
Best Local Similarity 41.6%; Pred. No. 3.1e-74;
Matches 266; Conservative 117; Mismatches 201; Indels 56; Gaps 13;

Qy 4 NAIRALFIQMYRRHQARRMQRCKNQIFONTLEYASEDOAE-----LYKFENDLIK 57
Db 90 STIKSALILQKMYRRCERLARARRATQIFALAEAGEOQLKRYVITLYDFADVIR 149
Qy 58 HMPQAGKKNQYQ--SAHVSVLDD--KDLVEEFGDIYNAKIF----- 97
Db 150 AMAENRGQGVNGNSPLMSALSHYAKPSLMSRGFTVKMLLEDTPTNVDIDRNYKGP 209
Qy 98 ---LPIRKNHIDILIDVFRKKRGRLAPKYVALILREASLSKQLPMSIPVSTAVSQOVT 154
Db 210 TLTSLPDRPQVAKMTEAF--KVNKVLHPKYVLMILHEARKIFKAMPVSRSRISTISNQVT 267
Qy 155 VCGDLHGRKLDLVLVLRKNGILPSSSNPYVFNQDFVDRGKGLVYLILLSLYLAFPAVFLNR 214
Db 268 ICGDLHGRKFDLCLITLVKNGIPSVNPIYIFNQGDFVDRGGGSIIVLCVLPALVIVDPMSIY 327
```

[illegible]

```

RESULT 4
T14614
hypothetical protein - Trypanosoma cruzi
C:Species: Trypanosoma cruzi
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14614
R:Anderson, B.; Aalund, L.; Pettersson, U.
submitted to the EMBL Data Library, March 1998
A:Description: 93.4 kb of complete sequence from chromosome 3 of Trypanosoma cruzi.
A:Reference number: Z18159
A:Accession: T14614
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-704 <AND>
A:Cross-references: UNIPARC:UPI000017B5B6; EMBL:AF052852; NID:g3063540; PTD:g3063551; PIR:
C:Genetics:
A:Map position: 3

```

```

Query Match 17.2%; Score 593.5; DB 2; Length 704;
Best Local Similarity 33.0%; Pred. No. 1.5e-29;
Matches 173; Conservative 73; Mismatches 204; Indels 75; Gaps 15;

QY 128 ILREAAKSLKOLPNI---SEPVSTAVS-----QQVVCGLDGLKLDLLVLRHKLGP 176
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 174 VFTDMSHLNTMPNVVRLSPPEVGARVSNGRINQSKVYVVDGLDGLADLLHLKEGMR 233
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 177 SSSNPYVYNGPVDGRKKGLVYLILLLSLYLAFPNVAFLNRGNHDSVMNARYCFIREVE 236
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 234 NEGYYIYINQGFVIRGANGVEVLILFSLMLACPCYVTLNRGNHRCYMDNDEYGFDEVS 293
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 237 SKYPRNHRILAFIDEVYRMLPLGSVLSNRVLVYHGF--SDSTSLIKSIDGKVVSI 294
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 294 TKYDRN---VRLVQRCCLAPLTIITIGKKYFVNHGGLPRKKGVNIEDISLQAFROI-- 348
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 295 LRPLPLDGEPLDKTKEWOQIF--DIMSDPQATMGCVPTLTKGAGVFGEDVTNDFLQHRRL 353
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 349 ---PMPD---YSQPEDEEIPQDLTMSDPEVDLQGMRSBPRGAGVFGADVTQBEFLQNNGL 402
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 354 SVVISHCECKPNGHFMFDNKITITIPASANTYVALGSNKGAY-ILANQOLMHPYQVYSAA 412
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 403 ELVIRSHCECLRGYEHHDKLLITVFSASNTDGPETNFGSAVAVVGDNPBESFTTYQAL 462
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 413 SQT-----RLSFQRMGIIVESSA-----LKELAVRNRHDELEL 448
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db      463 DDVEVQTLVDLGETPTPLGRVSSPATLSQSULLRRRRADVLAVLRRIRYORRRLLA 522

Qy -    449 BFRKYDPKDSGYISISHMKVMENYTKGLPWRLLDKLAPGDSQKAVNRTL----- 5020
        523 YFAKIDRTKRGSVMTKEWVEARVNTLNDLPMPFLRGVLVADDENTRVWVSHFLVGFHNP 5822

Qy      503 --DLDDTVILEAARDGMSVMDALYANKASLVAIRIIDLADNSGELTLDDEFETAILDLVA 560
        583 FOPLNTLNTWIOSA---CHHLLTQOORANHRQY---VAKAFNKEQVSVYEFCSVIRAIQY 6350

Db      561 HMPGAYSKAEMLEKCRMDLNDGDKVYDINRF-----LEAFRLSD 599
        636 TM----SEKQLFQLFVYLDGEGTGIIIDGKFTVNMLSMAATFLSD 676

```

RESULT 5
A55346
phosphoprotein phosphatase (EC 3.1.3.16) PPT [validated] - rat
N.Alternate names: serine/threonine phosphatase PPS
C.Species: Rattus norvegicus (Norway rat)
C.Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 28-Apr-2003
C.Accession: A55346
R.Becker, W.; Kentrup, H.; Klump, S.; Schultz, J.B.; Joost, H.G.
J. Biol. Chem. 269, 22586-22592, 1994
A.Title: Molecular cloning of a protein serine/threonine phosphatase containing a putative
A.Reference number: A55346; MUID:94357899; PMID:8077208

A: Molecule type mRNA
A: Residues: 1-499 <BRC>
A: Cross-references: UNIPARC:UP100001172827; GB:X177237
A: Note: authors translated the codon AAG for residue 53 as Gln, and GTA for residue 496
C: Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phosphatase
C: Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc
F: 28-61/Domain: tetratricopeptide repeat homology <T1>
F: 62-95/Domain: tetratricopeptide repeat homology <T2>
F: 96-129/Domain: tetratricopeptide repeat homology <T3>
F: 204-468/Domain: phosphoprotein phosphatase homology <P>
F: 236-305/Domain: phosphoesterase core homology <P>
F: 242,244,271/Binding site: iron (Asp, His, Asp) #status predicted
F: 271,303,352,427/Binding site: zinc (Asp, Asn, His, His) #status predicted
F: 272,304,451/Active site: Asp, His, Tyr #status predicted
F: 275,400/Binding site: substrate phosphate (Arg) #status predicted

```

Query Match 16.5%; Score 569; DB 1; Length 499;
Best Local Similarity 33.7%; Pred. No. 3.3e-28;
Match 143; Conservative 73; Mismatch 140; Indels 68; Gaps 16;

QY 11 FIOKRYRHQARRRORRCNMQIFONLEYSBEOQAEIYKFEFNDLIKMPQAGKXQYQ 70
   ::::: ::: ::: ::: ::: :::
DB 95 YIKGYTRRAASNMALGK-----FR---ALYRD---YEVVKKPKDKDAKKQYQ 137
   ::::: ::: ::: ::: ::: :::

QY 71 -----GSAR--VSVLDDKD-----DLVSEFG--DIVNAKIEIPIRKCHIDL 107
   ::: ::: ::: ::: ::: :::
DB 138 ECKSIYKQKAPERAIADGHRHSYVDSLIDIESMTIIEBYSQPKLEDGAVITITPKMD----- 193
   ::: ::: ::: ::: ::: :::

QY 108 LIDVFRKKRGRLHPKYYALLIREAASLQKLPNISPVSTAVSQVTVCGDLHGKLDL 167
   ::: ::: ::: ::: ::: :::
DB 194 LMQWYKQKQ--KLHRRKCAVQILVQKVELCKSLTEVETTLKETERKITVCGDTHGQFDLL 251
   ::: ::: ::: ::: ::: :::

QY 168 VILHKNGLPSSSNYYVNRGDPVDGKGLBYLLILLSLYLAFPNNAVFLNRGNHSDVWNA 227
   ::: ::: ::: ::: ::: :::
DB 252 NIFRINGLPSESTNPEYIFNGDFVDRGSPSVBEYILLTFGFKLLYPHFIHLRGNHEDNNMQ 311
   ::: ::: ::: ::: ::: :::

QY 228 RYGFIRLESYKYPNNHKKILAFIDEVYRWLPLGSVLSNRVLIYVGGF--SDSTSLDIKS 285
   ::: ::: ::: ::: ::: :::
DB 312 IYGBGBRYKAKYT---AQMTELFSBVEFWPLAOCINGKVLIMHGGLSEBGGVTLDDLRK 368
   ::: ::: ::: ::: ::: :::

QY 286 IDRGKYSILRPLTDEGFIDKTEWQOQIFDIIMSDDPOLATMGCVNPTLRGAGWFGPDTYD 345
   ::: ::: ::: ::: ::: :::
DB 369 IERRR-----QPF--DSGPM-----CDLWMSDPQPNQRSYSK--RGVSCQFGPDTYK 412
   ::: ::: ::: ::: ::: :::

QY 346 NPLQHRILSYVIRSHCEKPNGBHEPMDNKIIITFSASNYAIGSNKAYIRL--NQQLMPH 404
   ::: ::: ::: ::: ::: :::

```

Db 413 ARLSENQDLIIRSHVKAEGYVAHGRCVTVFSAFNYCDQGNKASYIHLQSGDLRQ 472

Qy 405 FVOY 408

Db 473 FHOY 476

RESULT 6

SS2570

Phosphoprotein phosphatase (EC 3.1.3.16) 5 [validated] - human

N/Alternate names: serine/threonine phosphatase pps

C/Species: Homo sapiens (man)

C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004

C/Accession: SS2570; PC4136

R/Chen, M.X.; McPartlin, A.E.; Brown, L.; Chen, Y.H.; Barker, H.M.; Cohen, P.T.W. EMBO J. 13, 4278-4290, 1994

A/Title: A novel human protein serine/threonine phosphatase, which possesses four tetrat

A/Reference number: SS2570; MUID:95009929; PMID:7925273

A/Accession: SS2570

A/Molecule type: mRNA

A/Residues: 7-498 <CH2>

A/Cross-references: UNIPROT:P53041, UNIPARC:UPI0000172826; EMBL:S73586; EMBL:X89416

A/Experimental source: teratocarcinoma cell line NTera-2

R/Xu, X.L.; Lagercrantz, U.; Zickert, P.; Bajalica-Lagercrantz, S.; Zetterberg, A. Biochem. Biophys. Res. Commun. 218, 514-517, 1996

A/Title: Chromosomal localization and 5' sequence of the human protein serine/threonine

A/Reference number: PC4136; MUID:96144708; PMID:8561788

A/Accession: PC4136

A/Molecule type: mRNA

A/Residues: 1-37 <XX>

A/Cross-references: UNIPARC:UPI000016B012; EMBL:X92121; NID:g1177477; PIDN:CA63089.1; F

A/Experimental source: fetal brain

C/Comment: mRNA encoding this protein was detected in all human tissues tested and was 1

A/Genetics: GDB:PPSC; PPS; PS

A/Genes: GDB:136857

A/Cross-references: GDB:136857

A/Map position: 19q13.3-19q13.3

C/Function:

A/Description: catalyzes the hydrolytic dephosphorylation of protein-phosphoserine [vali

A/Note: may play a role in the regulation of RNA synthesis and mitosis

C/Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phos

C/Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc

F/28-61/Domain: tetratricopeptide repeat homology <TT1>

F/62-95/Domain: tetratricopeptide repeat homology <TT2>

F/96-129/Domain: tetratricopeptide repeat homology <TT3>

F/1204-467/Domain: phosphoprotein phosphatase homology <PP>

F/1236-305/Domain: phosphoesterase core homology <PEC>

F/242.244,271/Binding site: iron (Asp, His, Asp) #status predicted

F/271.303,352,426/Binding site: zinc (Asp, Asn, His, His) #status predicted

F/274.304,450/Active site: Asp, His, Tyr #status predicted

F/273,399/Binding site: substrate phosphate (Arg) #status predicted

Query Match 16.5%; Score 567.5; DB 1; Length 498;

Best Local Similarity 34.0%; Pred. No. 4.1e-28;

Matches 144; Conservative 70; Mismatches 142; Indels 67; Gaps 16;

Qy 11 FIOKTRRQARREMRRCNMOIFONLBYASBODQALYKFFNDLIKHPQAGRNQOY 70

Db 95 YIKGYRRAASWALOK-----FR-----AALRD-----YEVVKKVPHDKAKKKQ 137

Qy 71 -----GSAHV-----SVLDDKD-----DVEEFG--DIYAKIELDIRKNHDL 107

Db 138 ECKKIYKQAFERAIAGDEHKASVSDLSIESKTIIDETSGPGLDEKVTISFMKE---- 193

Qy 108 LIDVPRKKRGNRLAPKRYVALILREAKSLKOLPNISPVSTAVSQOYVCGDLHGKLDL 167

Db 194 LMQWYDQK--KLHRCAGVQILVQVKEVLSKSTIVETTLKTEKIKTVGSDHGGFYDL 251

Qy 168 VILHKGGLPSSSNPYFVNDGDRGKGLSLVLLLSLYLAPNAVFLNRGNHDSVMA 227

Db 252 NIFELNGLSEETNPYLFENDFVDRGSGFSEVILITLGFGLLYPDHFLHKGHEETDNMQ 311

Qy 228 RGFIREVBSKYPNNKRLTAFIDVYRNLPLGSLVNSVLLVHGFSNST--SLDLISKI 286

Db 312 IYFEGEYKAKYT--AQMTLFSSEVFEWPLAOCINKKVLIMHGLSVKTYVLDLIDIKI 368

Qy 287 DRGKYVSLRPLTDEBPLDKTEWQOIFDIMSDPOATMGCVPTLRGAGYVFGPDVTN 346

Db 369 ERNR-----QPP--DSGPM-----CDLMSDPQONG--RSJBRGVSGQFGFDYTKA 412

Qy 347 FLQRRRLSYVRSHCKNGHEFMHNDKLTITTFASNTYVAGSKGAYIRL--NNQLMHP 405

Db 413 FLERNNDLYIIRSHVKAEGYVAHGRCVTVFSAFNYCDQGNKASYIHLQSGDLRQ 472

Qy 406 VQY 408

Db 473 HOY 475

RESULT 7

T45058

Phosphoprotein phosphatase (EC 3.1.3.16) Y398B ff [similarity] - Caenorhabditis elegans

N/Alternate names: serine/threonine phosphatase pps homology

C/Species: Caenorhabditis elegans

C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 28-Apr-2003

C/Accession: T45058

R/Wilson, R.; Ainscough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, J.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifkin, D.; Roopra, A.; Saunders, D. Nature 368, 32-38, 1994

A/Authors: Shownkeen, R.; Sims, M.; Smalton, N.; Smith, A.; Smith, M.; Sonhammer, E.; S

lock, L.; Wilkinson-Sprat, J.; Wohldman, P.

A/Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.

A/Reference number: 643531; MUID:94150718; PMID:7906398

A/Accession: T45058

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-526 <MTL>

A/Cross-references: UNIPARC:UPI0000083436; EMBL:AL132896; NID:96444440; PIDN:CA60937.1;

A/Experimental source: clone Y398B

C/Genetics:

A/Map position: 3

A/Introm: 42/1; 124/3; 184/3; 265/3; 348/3; 385/2; 415/2; 499/3

A/Note: Y398B. ff

C/Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phos

C/Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc

F/29-61/Domain: tetratricopeptide repeat homology <TT1>

F/62-95/Domain: tetratricopeptide repeat homology <TT2>

F/96-129/Domain: tetratricopeptide repeat homology <TT3>

F/203-497/Domain: phosphoprotein phosphatase homology <PP>

F/235-304/Domain: phosphoesterase core homology <PEC>

F/241.243,270/Binding site: iron (Asp, His, Asp) #status predicted

F/270,302,351,456/Binding site: zinc (Asp, Asn, His, His) #status predicted

F/273,303,460/Active site: Asp, His, Tyr #status predicted

F/274,429/Binding site: substrate phosphate (Arg) #status predicted

Query Match 16.2%; Score 558; DB 1; Length 526;

Best Local Similarity 31.6%; Pred. No. 1.7e-27;

Matches 138; Conservative 79; Mismatches 172; Indels 48; Gaps 14;

Qy 3 ENAIRPAI---FIOKTRRQARREMR--RCNMOIFONLBYASBODQALYKFF--FNDLI 56

Db 86 DNAL--ALDPYVVGFGYRATANNALGRFKKALTDYQVAVVYCPNDXARAKFDBCSKIV 143

Qy 57 KMPQAAKRNQOYQGSANVSVDKDDLVSEFGDIYNAKILPFRKNHIDLIDVFRKR 116

Db 144 RROKFEAISTDHDKQVAETLIDNAMAIBESYD--GRLDCKITKEFVLQILKIFKQO 201

Qy 117 GNRLLPKRYVALILREAAKSLKOLPNISPVSTAVSQOYVCGDLHGKLDLVLVHKNGLP 176

Db 202 --KLHKKTAFFKMLEFYTVKSLPFWETIYPTGKKFTICGVHQGFYDLCKHIFINQY 259

Qy 177 SSSNPYVNDGFDVDRGKGLSLVLLLSLYLAPNAVFLNRGNHDSVMAVYGFIRVE 236

Db 260 SETNPYLVNDGFDVDRGSGSVETIFPMIGFKLLYPRHFMKSGNHSDVMNMMYGBEGVK 319

| | | | |
|----|-----|---|-----|
| Qy | 237 | SKYENHRRILAFIEVVRMLPLGSVLSRVLIVYGG--SDSTLIDIKSIDGKXYSI | 294 |
| Db | 320 | AKYT---QQCMDFETFCWLPPLCHLINEKLIIVCHGGLFKBEDVTLIEDIRTDNR---- | 371 |
| Qy | 295 | LRLPLTDEPLD---KTEMQOIFDIM-----SDPQATMGCVPTLL | 332 |
| Db | 372 | -RQPRDEGIMCDLLEKKNKQUL-KILYPDGKINKNNKSNQPKTCNADSPQINRSPSK- | 428 |
| Qy | 333 | RGAGVWFEPDYTDNFLQRRRLSYVIRSHCEKRNHGFMDNKKLITIPASNYVAIGSNKG | 392 |
| Db | 429 | RGVGQGFQBPDYTSKCEKTRTGIEVYVRSHEVPRDEGEHMHNGCJFTVFSAPNYCDQMNKG | 488 |
| Qy | 393 | AYIRL--NQQLMHPFVQY 408 | |
| Db | 489 | AFITITGDLTPRTFP 505 | |

RESULT 8
E84858
phosphoprotein phosphatase (EC 3.1.3.16) At2g42810 [similarity] - Arabidopsis thaliana
C|Species: Arabidopsis thaliana (mouse-ear cress)
C|Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 28-Apr-2003
C|Accession: E84858
R|Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.B.; Unayam, L.; Tallon, L.;
Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, A.
Nature 402, 761-766, 1999
A|Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A|Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: B94858
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-533 <STO>
A:Cross-references: UNIPARC:UPI0000044D6A; GB:AE02093; NID:g4512673; PIDD:AA2127.1; C:GeneticCB:
A:Gene: At2g42810
A:Map position: 2
C:Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phospho
C:Keywords: phosphoric monoester hydrolase

```

Query March 16.2%; Score 557.5; DB 2; Length 533;
Best Local Similarity 39.4%; Pred. No. 1.9e-27;
Match 127; Conservative 55; Mismatches 109; Indels 31; Gaps 10;

Qy 94 AKIE-LPRKHHIDLLIDVFRKKGGRNLAPKQVALILIREAKSLKOLEPNISPVTAASQ 152
   ||| : : : : : ||| : : : : :
Db 214 ARISEBETLDPVKTMEDPFKKQK--TLHKRAYQIVLQTIQILALPSLVDISVPHKX 271
   ||| : : : : : ||| : : : : :
Qy 153 VTVCGDLHGKLDLLVLHKNGLPSSSNPYVFNQDPVDGRGRGLEVLILLSLYLAPEA 212
   ||| : : : : : ||| : : : : :
Db 272 ITVCGDVHGFYDLINLIFELNGLPBENPVLFNQDPVDGRGSFVAILITLPAFCMCSS 331
   ||| : : : : : ||| : : : : :
Qy 213 VFLARGNHEDSVNARVGYPIREVBESKYPFRNHKRIILAFID---EYRWLPCLGSVLNSRYLI 269
   ||| : : : : : ||| : : : : :
Db 332 IYLARGNHESKSMNKTYGFEGBVRSKLSK-----FVDLEAFVFCYLPFLAHVINGKQFV 385
   ||| : : : : : ||| : : : : :
Qy 270 VHGG-FS-DSTSLDIKSIDRGKYVILRPPLTIDGSRPLDKTIEMQOIPIMSSDPOATWVC 327
   ||| : : : : : ||| : : : : :
Db 386 VHGGILFSVDGKYLSTIRALDR-----PCSP-----BGLMCLMSDPOPLFGR 430
   ||| : : : : : ||| : : : : :
Qy 328 VENTLRGAGVWFPGPDVDNPLQRLHLASYIRSHCEKPGHGFEMDNKIITTFASANYAI 387
   ||| : : : : : ||| : : : : :
Db 431 GPSK-RGVGLSFEGGVDVTRPLQDNVLLDLVSHSEKDEGVREYHDGKLIITVFASPNYCDQ 489
   ||| : : : : : ||| : : : : :
Qy 388 GSNKGAVYRLN-NQLMPEFYQ 408
   ||| : : : : : ||| : : : : :
Db 490 MGNKGAFIRFEAPDKKPNIVTF 511
   ||| : : : : : ||| : : : : :

```

RESULT 9
T46576
phosphoprotein phosphatase (EC 3.1.3.16) ptc-1 [similarity] - *Neurospora crassa*
N:Alternate names: serine/threonine phosphatase PPS homolog

C:Species: *Neurospora crassa*
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 28-Apr-2003
C:Accession: T46576
R:Yatzen, E., Yarden, O.
submitted to the EMBL Data Library, February 1997
A:Description: Pct-1, a N. crassa novel-type phosphatase.
Reference number: Z23089

A/Accession: T46576
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-479 <YAT>

A:Gene: ppc-1
A:Map position: V
C:Superfamily: protein phosphatase 5, phosphoesterase core homology; phosphoprotein phosphatase
C:Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc
F:8-41/Domain: tetratricopeptide repeat homology <TT1>
F:42-75/Domain: tetratricopeptide repeat homology <TT2>
F:76-109/Domain: tetratricopeptide repeat homology <TT3>
F:183-448/Domain: phosphoprotein phosphatase homology <PPP>
F:215-284/Domain: phosphoesterase core homology <PBC>
F:221, 223, 250/Binding site: iron (Asp, His, Asp) #status predicted
F:250, 282, 331, 408/Binding site: zinc (Asp, Asn, His, His) #status predicted
F:253, 283, 432/Active site: Asp, His, Tyr #status predicted
F:254, 351/Binding site: substrate phosphate (Arg) #status predicted

| | | | | |
|-----------------------|------------------|-----------------|------------|-------------|
| Query Match | 15.7% | Score 541; | DB 1; | Length 479; |
| Best Local Similarity | 32.3% | Pred. No. 1 | 8e-26; | |
| Matches 140; | Conservative 57; | Mismatches 145; | Indels 88; | Gaps 10; |

```

Qy 11 FIOQMYSRHOQARREMOBCRNWOJFONLEVASBODQAEIYKFFNDLIKHMPOAARGKQOY 70
   |:::|::|
Db 75 FVKAYYRAT- - - - - YAIIMPKEAVKOPKTCVKLA- - - - - 107

Qy 71 GSAHVSUDDQD- - - - - DLYE- - - - - BPGD.VNAKIELPIRKNHIDLI- - - 110
   |:::|::|
Db 108 - - - - - DNMDAKLKIVCEKCIWROLAPFAAIEVGBELISAABGIDVESMAVDASDYGR 159

Qy 111 - - - - - VPRKRGNRLLHPKVALLIREBAKSLQOLPNISVSTAVSQOVTYC 156
   |:::|::|
Db 160 LEGHEMTQEPFDIMIERPFKRGKLIHKYVYOTIIIAVRNIVNEPRMWEVDIPEDEVQLYVC 219

Qy 157 GDHLHGKLDLILVLAHKQGLPSSSNPYFNGDFVDRGKKEVLILLLSLYLAFNPAVPLN 216
   |:::|::|
Db 220 GDTHGQYFDMLELFRLLNGPSPDKHYLLFENGPFVDRGSMSTELALLVAYKMLRPRGFFIN 279

Qy 217 RGNHEDSVNARAYGFIRBEVBSKYPRNHRKILAFIDEVRYMLPLGSLVNSRVLIVHGSP- - 274
   |:::|::|
Db 280 RGNHRETDMMNRVYVFEGBECKHY- - - NERTYKLFSESPSALPLTILIKKFLVHLHGGLFS 336

Qy 275 SDSTSLLISIDBGKYVSLIRPPLTGEPLDKTEWQOIFIMMSDQATNGCVNLTLRG 334
   |:::|::|
Db 337 DDNVTLLDIRKLDLHKO- - - KQPGQG- - - - - LMMEMLMTDPPPPRGRGSP- RG 382

Qy 335 AGVWFGPDVTNPLQRRHLSTVIRSHCEKPGNGBHEHMDNKLIITFSASNYAIGSKNGY 394
   |:::|::|
Db 383 VGMQFGPDVTNRFPDCKKGLBAIRSHVRYMDGYEBEHGKCIYVSAFYCDMTENKRAY 442

Qy 395 IRLANQMLRPHVOY 408
   |:::|::|
Db 443 INIGPDYKLFKSOP 456

```

RESULT 10
 T40391
 phosphoprotein phosphatase (BC 3.1.3.16) SPBC3P6.0c [similarity] - fission yeast (Schizosaccharomyces pombe)
 N:Alternate names: aserin/threonine phosphatase pP5 homolog
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 28-Apr-2003
 C:Accession: T40391
 R:Lyne, M.; Kalandream, M.A.; Barrel, B.G.; Churche, C.M.
 submitted to the EMBL Data Library, February 1998

A:Reference number: Z21925
A:Accession: T40391
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-473 <LYN>
A:Cross-references: UNIPARC:UPI000006B641; EMBL:AL022019; PIDD:CAAL7690.2; GSPDB:GN0006
A:Experimental source: strain 972h-; cosmid c3f6
A:Genetic(s):
A:Gene: SPDB:SPBC396.01C
A:Map position: 2
C:Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phospho
C:Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc
F:5-38/Domain: tetratricopeptide repeat homology <TT1>
F:39-72/Domain: tetratricopeptide repeat homology <TT2>
F:73-106/Domain: tetratricopeptide repeat homology <TT3>
F:119-444/Domain: phosphoprotein phosphatase homology <PPP>
F:211-280/Domain: phosphoesterase core homology <PEC>
F:217,219,246/Binding site: iron (Asp, His, Asp) #status predicted
F:246,278,327,404/Binding site: zinc (Asp, Asn, His, His) #status predicted
F:249,279,428/Active site: Asp, His, Tyr #status predicted
F:250,377/Binding site: substrate phosphate (Arg) #status predicted

Query Match 15.1% Score 519; DB 1; Length 473;
Best Local Similarity 35.2%; Pred. No. 4,3e-25;
Matches 119; Conservative 66; Mismatches 127; Indels 26; Gaps 8;

QY 73 AHVSLDTRKDDLVEFGDVIYNAKIELPIRKHIDLLIDVFAKKGNRLHPKYVALIREA 132
DB 139 ANINIED--MIPSPYDGI---LEKQITKFEVDMERF---CQKULPKFAYSIIRD 191
QY 133 AKSLKQLEPNISFVSTAVSQQVTVCGDLHGKLDLIVLHKGLDPSSSNPVYFNGDPVDRG 192
DB 192 KELLEKTPSLIDIPYKQDETLVTCDDTHQGFYDLNLTKLHGPPSPPTKYLFGNDPVDRG 251
QY 193 KRGLVLELLLSLYLAFNNAVFLNRGNHEDSVNARVCFITFVSKYPRNKRILATIDE 252
DB 252 SWSTEVATTLVAYKLLIFDPAVFINRGNHETDDNNVYGFSECSRSKY---NERTFNIFSE 308
QY 253 VYRWLPFGSVLNSRYLVHGGF--SDSTSLDLIKSIDRGKYSILRPLTDGEPLDXTW 310
DB 309 TFSLLPLGSLTSDSYLVVHGGFLPSDDNTVTLQNRIDFS---KKQPGQSG----- 356
QY 311 QQIFPIMMSDPQATMGCVPTLRGAGVFPFGDVTNDFLQRRRLSYVIRSHCKENGHEFM 370
DB 357 -LMEMLMTDTPQAPGRGPRSK-RGVGLQFGPDVSRRFCEANGLAVIRSHSEVRDQGYEVE 414
QY 371 HDNKLTITFASANSYVIAISNGKAYVIRLNNOLMHPHYQY 408
DB 415 HDGYCTIVFSAPNYCDSTGNLDAVIVKREMDDELDPHP 452

RESULT 11
S52571
phosphoprotein phosphatase (EC 3.1.3.16) PPT1 - yeast (*Saccharomyces cerevisiae*)
N/Alternate names: protein G6347; protein YGR123C; serine/threonine phosphatase PPS homoc
C/Species: *Saccharomyces cerevisiae*
C/Date: 03-Aug-2001 #sequence revision 03-Aug-2001 #text_change 09-Jul-2004
C/Accession: S52571; S55981; S64432; S64697
R/Chen, M.X.; McPartlin, A.E.; Brown, L.; Chen, Y.H.; Barker, H.M.; Cohen, P.T.W.
EMBO J. 13, 4278-4290, 1994
A/Title: A novel human protein serine/threonine phosphatase, which possesses four tetrat
A/Reference number: S52570; MUID:J95009929; PMID:7925273
A:Accession: S52571
A:Molecule type: mRNA
A:Residues: 1-513 <CHR>
A:Cross-references: UNIPROT:P50343; UNIPARC:UPI0000053298
R/van Dyck, L.; Goffeau, A.
submitted to the EMBL Data Library, December 1994
e new ORFs, remnants of Ty and three tRNA genes.
A:Reference number: S55976
A:Accession: S55981
A:Molecule type: DNA

A:Residues: 1-53 <VAM>
A:Cross-references: UNIPARC:UPI0000053296; EMBL:X83099; NID:g6423340; PID:g6423346
R:Van Dyck, L.; Skala, J.; de Weydts, P.; Purnelle, B.; Talla, E.; Nawrocki, A.; Del
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64428
A:Accession: S64432
A:Molecule type: DNA
A:Residues: 1-513 <VAM>
A:Cross-references: UNIPARC:UPI0000053296; EMBL:Z72908; NID:g1323200; PID:e243462; PID:g
A:Experimental source: strain S286C
R:Chen, M.X.; McPartlin, A.E.; Brown, L.; Chen, Y.H.; Barker, H.M.; Cohen, P.T.W.
submitted to the EMBL Data Library, July 1995
A:Description: A novel human protein serine/threonine phosphatase, which possesses four
A:Reference number: S64697
A:Accession: S64697
A:Molecule type: DNA
A:Residues: 1-380, 'HMLR', 369, 'AP', 392-393, 'NFGRTHKPKPGVVL', 409, 'SVV', <CHW>
A:Cross-references: UNIPARC:UPI00001668B1; EMBL:X69417; NID:g897805; PID:g897806
A:Genetic: S:
A:Gene: SGD:PPT1
A:Cross-references: SGD:S0003355; MIPS:YGR123C
A:Map position: 7R
C:Superfamily: protein phosphatase 5; phosphatase core homology; phosphoprotein phosph
C:Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc
F:12-45/Domain: tetratricopeptide repeat homology <T1>
F:46-79/Domain: tetratricopeptide repeat homology <T2>
F:80-113/Domain: tetratricopeptide repeat homology <T3>
F:209-475/Domain: phosphoprotein phosphatase homology <P1>
F:243-312/Domain: phosphatase core homology <P2>
F:243-251,278/Binding site: iron (Asp, His, Asp) #status predicted
F:278,310,359,434/Binding site: zinc (Asp, Asn, His, His) #status predicted
F:282,311,458/Active site: Asp, His, Tyr #status predicted
F:282,407/Binding site: substrate phosphate (Arg) #status predicted

Query Match 13.9%; Score 479.5; DB 1; Length 513;
Best Local Similarity 32.3%; Pred. No. 1.4e-22;
Matches 130; Conservative 66; Mismatches 154; Indels 51; Gaps 13;

2 DENAIRAIRIFOKYRRHQAARRM--QRRCNNQIFONTLYASRQDAQLYKEFNDLIKH 58
113 DPATKALLVCDRIRERFRKALIGAEENAKISLQCTLNLSFDNADLAN----- 165
59 MPQAGKKNQYQSSAHVSLDDKDLVBFSGDIVNAKISLPIRKNIHLLIDVFRKKGN 118
166 -----EGKTLKFE-----QLYDDKK-----AFKGAKKIMQGEFISKWVNDLFLK--GK 207
119 RLHKYVALILREAKSLKQLPNISPY--STAVSQVTVGGDLHGKLDLLVYLHNGLP 176
208 YLPKRYAAILISHADTLFRQPSNVELENNSTDVAKISVCGDTHGGPYDVLNFRFGKY 267
177 SSSNPYVFNDDFYDRGKGLVLLLLSLTSLAEPNVAFLRGHSDSVNNAVGYFIREVE 236
268 GKPTITVFNDDFYDRGMSCEVALLPCLKILAPNNFPLNRGNHSDNNMKITGYFEDCK 327
237 SKYPRNKRKILAFIDEYRMLPLGSVLSKVLVHGQF--SD--STSLDLKISIDRGKYSI 294
328 YKY---SQRIFNMPAGFESLPLATILNNDYLVHGGLPESDPATILSDFNKIDR-----F 379
295 LRPLTLGSEPLDKTEWQITDIMSDDPQATMGCVPTNLRGAGVFRGPDVTNLTQHRRLS 354
380 AQPP-RDG-----AFNELMADPOEANGKGPSP--RGLGHAFPDIDTRFLRNKLR 428
355 YVIRSHCKDNGHGFHMDNKIITPSASNYVAIGSNKGAIVRL 397
429 KIFRSHLRMGVGFQEGKGLMTVFSAPNYCDSQGLDGVIHV 471

RESULT 12
phosphoprotein phosphatase (EC 3.1.3.16) PF7 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C:Accession: T51611

submitted to the EMBL Data Library, July 1996

A:Reference number: Z19402

A:Accession: T21288

A:Status: translated from GB/EMBL/DDBB

A:Molecule type: DNA

A:Residues: 1-401 <W/L>

A:Cross-references: UNIPROT:Q27495; UNIPARC:UPI000007C7BF; EMBL:Z77659; PIDD:CAB01164.1

A:Experimental source: clone F23B12

C:Genetics:

A:Gene: CESP:F23B12.1

A:Map position: 5

A:Introns: 71/3; 121/3; 153/2; 251/2; 377/2

C:Superfamily: serine/threonine protein phosphatase, phosphoesterase core homology, phosphoric monoester hydrolase, serine/threonine-specific kinase

R:197-356/DNA: phosphoprotein phosphatase homology <PP>

R:125-193/DNA: phosphoesterase core homology <PEC>

F:131,133,159/Binding site: iron (Asp, His, Asp) #status predicted

F:159,191,240,315/Binding site: zinc (Asp, Asn, His, His) #status predicted

F:162,192,333/Active site: Asp, His, Tyr #status predicted

F:163,288/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.3%; Score 425; DB 1; Length 401;
Best Local Similarity 29.3%; Pred. No. 2,7e-19;
Matches 113; Conservative 66; Mismatches 137; Indels 68; Gaps 12;

OY 50 KFPNDL-----IKHPQAARKKQYOGSAHVLYDDKDVLVEBFDIVNAXIEPIRKQ 103
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 13 KEFGDVFHVDLFIFRPLSLNONSIT---SSVSLT-SKSSMEVP-----ESSAISRN 63
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 104 HIDLIDIVRRKKRGKRNLHP-----KYALLIREAKSLKQLPNIS 143
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 64 AAG-NDVQKNKQNMHSVIERLKWSPGRCQQLFVENELIELCYRARDEPFWK-----N 116
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 144 PVSTAVSQGVTCGDLGLKDDLVLVAHKNGLPSSSNPVYVNGDPVDRGKGLLEYTLILL 203
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 117 KYKLIDIAEVKLTGGDIHGQPEDLMALPELNCP--BEHKTYLFLGDVYDGPSIEVITYLLF 175
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 204 SLYLAFPNVFLNRGNHSDVMNARYGFIREVESKYPNNRIILAFIDE---VYRWLPGLG 260
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 176 TFOIIMPDKVFTLRGNHESRPVNMQGYFLCKKRYTS-----VALYDAFOLAPNCMPIC 229
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 261 SYLVNSRVLLVHGFS-DSTGLDLIKSIDRGKTVSILRLPPLTDGEPLDTIEWQQIPDIMMS 319
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 230 AVVSKIIICMGHGISSELDLIDLTLEKIDR-----PFDIPDIGVISDLTWA 274
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 320 DPOATMGCVPNTIRGACGWFGPDVTDFLOHRISVYIRSHCKKNGHEPHMDNKIITTF 379
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 275 DDEKEFGVADSPRGASRSGPNVAKKFLDMHNHLVVRAHQVMDGYEFPADRQLVTVF 334
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 380 SASNYVAIGSNKGAYIRLNQQLMPHF 405
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 335 SAPSYCGQEDMAAANMVNDKLCCTF 360
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 23
catalytic chain (clone TOPP2) - Arabidopsis thaliana
phosphoprotein phosphatase (EC 3.1.3.16)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S31086
R:Smith, R.D.; Walker, J.C.
Plant Mol. Biol. 21, 307-316, 1993
A>Title: Expression of multiple type 1 phosphoprotein phosphatases in Arabidopsis thaliana
A:Reference number: S31085; MUID:J3144705; PMID:7678768
A:Accession: S31086
A:Molecule type: mRNA
A:Residues: 1-312 <SMI>
C:Superfamily: serine/threonine protein phosphatase, phosphoesterase core homology, phosphoric monoester hydrolase, serine/threonine-specific kinase
R:156-295/DNA: phosphoprotein phosphatase homology <PP>
R:164-132/DNA: phosphoesterase core homology <PEC>

F:710_72_98/Binding site: iron (Asp, His, Asp) #status predicted
F:710_110_119_254/Binding site: zinc (Asp, Asn, His, His) #status predicted
F:101_111_278/Active site: Asp, His, Tyr #status predicted
F:102_227/binding site: substrate phosphate (Arg) #status predicted

Query Match 12.3%; Score 42.1; DB 2; Length 312;
Best Local Similarity 30.5%; Pred. No. 2,1e-19;
Matches 109; Conservative 63; Mismatches 128; Indels 57; Gaps 10;

QY QVGSAAHVSLDDKDLVEBFDIVNAKILPFRKNHIDLLIDVPRKKRG-----NRLA 121
DB 4 QCGSMDPALDD-----IIR-----LIDYRNPCKGQAMLNSE 40
QY 122 PKYALLIREAKSKLPINISPVSTAVSQOVTWCGGLKGDLDDLVLHKNGLPSSNP 181
DB 41 IROLCIVSRRI---FLQOPMLLEAPRK-----ICGHIHQGYSDILLRLFVGSGPPPTAN- 92
QY 182 YVPNGDVPDGKRGLGVTLILLSLYLAPPNAVPLNRNHDSVNNARVGFREVEREKPR 241
DB 93 YLFGLADVVDKQSLETTCLLAYKIKYPENFFLRKNHCASINRYIGFDECKRR-- 150
QY 242 NHRKIILAFIDEVVRMELPLGSVLTNSRVLIIVHGFS--DSTSLDIKSIDRGKYVSIILRPPLT 300
DB 151 -SVALKMVFTDPSFCFLVAAVIIDDKLICMHGGLSPDLTNVGIKNIR----- 197
QY 301 DGEPLDKTEWQQIFDINMSDPQATMGCVPTLRGAGVPGPDVTDNFLQRRLSYIRSH 360
DB 198 ---PLTPDVDSGLCDLLMSPDSKOVKGMNDRGVSYTGPDKYAERLIKNDMLICRAH 254
QY 361 ECRKHGHFEHDNKKIIITFPASNYYAIGSKGAIYRLNQMLPHFVOYISAASQTKR 417
DB 255 QVEDVGEYEPADROLVTFPSAPNYCGSPDNAGAMSVDSELMCSF-QILKPADRKKPR 310

RESULT 24
T09550
phosphoprotein phosphatase (EC 3.1.3.16) 1, catalytic epsilon chain - alfalfa
C/Species: Medicago sativa (alfalfa)
C/Dates: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T09550
R/Vassi, K.; Georgas Toch, B.; Kovacs, L.; Magyar, Z.; Horvath, G.; Bagosai, P.; Gergely,
Arch. Biochem. Biophys. 360, 206-214, 1998
A>Title: Protein phosphatase 1 catalytic subunit isoforms from alfalfa: Biochemical char
A/Reference number: Z16730; MUID:99068922; PMID:9851832
A/Accession: T09550
A>Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-323 <UN>
A/Cross-references: UNIPROT:O65847; UNIPARC:UPI00000AABFC; EMBL:AJ002488; NID:g3176077;
A/Experimental source: subspecies sativa; strain RA3
C/Genetics:
A/Gene: PPI epsilon
C/Function:
C/Description: catalyzes hydrolysis of peptidyl-phosphoserine or -phosphothreonine to r
C/Superfamily: serine/threonine protein phosphatases; phosphoesterase core homology; pho
C/Keywords: phosphoric monoester hydrolase; serine/threonine-specific phosphatase
F:42-301/Domain: phosphoprotein phosphatase homology <PP>
F:710-138/Domain: phosphoesterase core homology <PEC>

Query Match 12.2%; Score 42.1; DB 2; Length 323;
Best Local Similarity 32.5%; Pred. No. 3,6e-19;
Matches 106; Conservative 58; Mismatches 132; Indels 30; Gaps 7;

QY 107 LLIDVPRKKGGRNLHPKYVALILIREAKSKLPNIS-----PSTAVSQOVTWGD 159
DB 20 VLDDIIRRLREVRISRPGKQLSEA--ETKOLCSARDIFLQOPNLBLEAPKICGDI 77
QY 160 HGKDDLDLVVLHKUGLPSSSNPVFNGDFVDRGKRGLEVLLLSLYLAEPNAVFLNRGN 219
DB 78 HGGYSDDLRLPEYGGULPQSN-YLFLGDYDRGQSLETTCLLAYKIKYPENFFLRGN 136
QY 220 HEDSVNARVGFIREVESKTPRNHRIILAFIDEYTRMLPLGSVINSLVLIIVHGFS--DST 278

Db 137 HECASINRIYGFYDECKRRF---NVALMKAFIESFNCFLVAAALIDSKILCMHGSLPDLT 193
Qy 279 SLDLISIDRGKVVSLRPPPLDGBELDTKTEWQOIFDIMSPOATMGCVPTLTGAGW 338
Db 194 NIDQJINLPR-----FVPIPDGILMDQFMSDKVCKVKCMQANMDGVSYT 238
Qy 339 FGPDVDFNFQIRHLSYVIRSHCEKNGHFMHNDKTIITFSASNYAIGSKAGYIRLN 398
Db 239 FGPDKAASFLTRNDLILCRHQVAVDGEFFPADROLVITFSAPNYCGEFNDAGAMMSVD 298
Qy 399 NQIMPHFVOYISAASQTKSLSPKQRM 424
Db 299 ENLMCSF-QILKPAKKTGFVMSNKM 323

RESULT 25
T09548
phosphoprotein phosphatase (EC 3.1.3.16) 1 catalytic chain delta - alfalfa
C:Species: Medicago sativa (alfalfa)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T09548
R:Vissel, E.; Cserotdae Toch, E.; Kovacs, I.; Magyar, Z.; Horvath, G.; Bagossi, P.; Gergely Arch. Biochem. Biophys. 360, 206-214, 1998
A:Title: Protein phosphatase 1 catalytic subunit isoforms from alfalfa: Biochemical char
A:Reference number: Z16730; PMID:99068922; PMID:9851832
A:Accession: T09548
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-326 <VIS>
A:Cross-references: UNIPROT:O65846; UNIPARC:UPI00000AAB92; EMBL:AJ002487; NID:G3176075;
A:Experimental source: subspecies sativa, strain RA3
C:Genetics:
A:Gene: PPI delta
C:Function:
A:Description: catalyzes hydrolysis of peptidyl-phosphoserine or -phosphothreonine to re
C:Superfamily: serine/threonine protein phosphatase; phosphoserase core homology; phos
C:Keywords: phosphoric monoester hydrolase; serine/threonine-specific phosphatase
F:26-285/Domain: phosphoprotein phosphatase homology <PPP>
F:54-122/Domain: phosphoserase core homology <PESC>

Query Match 12.2%; Score 421; DB 2; Length 326;
Best Local Similarity 30.9%; Pred. No. 3.6e-19;
Matches 104; Conservative 67; Mismatches 128; Indels 36; Gaps 9;

Qy 105 IDLLIVPFRKGNRLHPKVALILRBAKSLKQLENIS-----PVSTAVSQQVTVCG 157
Db 10 INRLLEV-RGRPEKQVQ-----LSBA--EIKQLCVSKDIFMNPMLIKLEAPIKICG 59
Qy 158 DLHGKLDLLVVLHKNGLPSSSNPYVFNQDGVDRGKGLVLLILSLYLAFPNAVFLNR 217
Db 60 DLHGQVSDLRLEFYGGFPPRRN-YLFLGDVYDRGKQSLTICLLAYIKIKIPKQFFLR 118
Qy 218 GNHEDSVNARVYGFIREVSKYPRNKRILAFIDVYRWPLGQSVLNSVLLVHGFS-D 276
Db 119 GNHECASINRIYGFYDECKRRY---NVKLMQWTFDQFNCFLPAALIDKILCMHGSLSP 175
Qy 277 STSLDLISIDRGKVVSLRPPPLDGBELDTKTEWQOIFDIMSPOATMGCVPTLTGAG 336
Db 176 LHDLMQIKMLRR-----PCEVPESGILLCDLMSDPSDVRGMSRSGVS 220
Qy 337 VWFPGVPTNFIQIRHLSYVIRSHCEKNGHFMHNDKTIITFSASNYAIGSKAGYIR 396
Db 221 YTFGADRVKVEFLQKHLDLILCRHQVAVDGEFFPAROLVITFSAPNYCGEFNDAGAMMT 280
Qy 397 LNNQIMPHFVOYISAASQTKSLSPKQRMGIVSSALK 433
Db 281 VNESIVCSF-QILKPDKPKKPSFGSTTYKSSSPK 316

RESULT 26
T03304
probable phosphoprotein phosphatase (EC 3.1.3.16) 1 catalytic chain - rice
C:Species: Oryza sativa (rice)

C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C:Accession: T03304
R:Wang, B.; Chang, M.; Chen, X.; Wu, R.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z14900
A:Accession: T03304
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-317 <MAN>
A:Cross-references: UNIPROT:P48489; UNIPARC:UPI0000131BPF; EMBL:U31773; NID:G951335; PID
A:Experimental source: seed
C:Superfamily: serine/threonine protein phosphatase; phosphoserase core homology; phos
C:Keywords: phosphoric monoester hydrolase
F:36-295/Domain: phosphoprotein phosphatase homology <PPP>
F:64-132/Domain: phosphoserase core homology <PESC>

Query Match 12.2%; Score 420.5; DB 2; Length 317;
Best Local Similarity 31.6%; Pred. No. 3.8e-19;
Matches 103; Conservative 61; Mismatches 123; Indels 39; Gaps 10;

Qy 108 LIDVFRKRGRLHPKVALILRBAKSLKQLENIS-----PVSTAVSQQVTVCGDLH 160
Db 22 LLEVTAAPGKQVH-----LSBS--EIKQLCVSRIFLSQPLILRLKAPIKICGDIH 72
Qy 161 GKLDLILVVLHKNGLPSSSNPYVFNQDGVDRGKGLVLLILSLYLAFPNAVFLNRGN 220
Db 73 GQVSDLLLFYGGFPPRRN-YLFLGDVYDRGKQSLTICLLGYKIKYIPENFLRNH 131
Qy 221 EDSVNNARVYGFIREVSKYPRNKRILAFIDVYRWPLGQSVLNSVLLVHGFS-DSTS 279
Db 132 HECASINRIYGFYDECKRRF---NVALMKAFIESFNCFLVAAALIDSKILCMHGSLPDLT 188
Qy 280 LIDLISIDRGKVVSLRPPPLDGBELDTKTEWQOIFDIMSPOATMGCVPTLTGAGW 338
Db 189 LDEIRKLAR-----PTDVPDTGLCDYFMSDVGKQVQGMQAND-RGVSYT 232
Qy 339 FGPDVDFNFQIRHLSYVIRSHCEKNGHFMHNDKTIITFSASNYAIGSKAGYIRLN 398
Db 233 FQADVSEFLRHGDLILCRHQVAVDGEFFPADROLVITFSAPNYCGEFNDAGAMMSVD 292
Qy 399 NQIMPHFVOYISAASQTKSLSPKQRM 424
Db 293 ENLMCSF-QILKPAKKTGFVMSNKM 317

RESULT 27
T21322
phosphoprotein phosphatase (EC 3.1.3.16) 1 P25B3.4 [similarity] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 05-Oct-2004
C:Accession: T21322
R:Gardner, A.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19406
A:Accession: T21322
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-291 <WII>
A:Cross-references: UNIPROT:Q27496; UNIPARC:UPI000007D589; EMBL:Z70752; PIDN:CAA94756.1;
A:Experimental source: clone F25B3
C:Genetics:
A:Gene: CESP:P25B3.4
A:Map position: 5
A:Anticodon: 20/2, 157/3, 273/3
C:Superfamily: serine/threonine protein phosphatase; phosphoserase core homology; phos
C:Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci
F:7-268/Domain: phosphoprotein phosphatase homology <PPP>
F:35-103/Domain: phosphoserase core homology <PESC>
F:41,43,69/Binding site: iron (Asp, His, Asp) #status predicted
F:69,101,151,227/Binding site: zinc (Asp, Asn, His, His) #status predicted
F:72,102,251/Active site: Asp, His, Tyr #status predicted
F:73,200/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.2%; Score 420, DB 1, Length 291;
Beet Local Similarity 35.2%; Pred. No. 3.6e-19;
Matches 94; Conservative 48; Mismatches 107; Indels 18; Gaps 6;

Oy 141 NISPVTAASQOATVCGDLHGKLDLVLHKNGLPSSSNPYFVGDFYDRGKLEVL 200
| :
Db 24 NKQPMILIEVNSPINICGDHIGQPSDRLRFPKNGPPHRAN-YLFAGDYVDKKKCLITLL 82
| :
Oy 201 LLSTLYAPPNAVFLNRGNHEDSVGMARYPGIRREVSKYPRNKRILLAFIDEVEMLPYG 260
| :
Db 83 LLEFAVKVIFPNHFMLRGNHCSLIINRQGYEEQRKY--NKPEVMHSFGCVSWPLT 140
| :
Oy 261 SYLVNSVYLIVHGFSDDSLDIKSIDRKVSYILRPPLDTGEPLDKTEWQOIPIIMMSD 320
| :
Db 141 ALVGGRIQLCHGGVS-----KMLQNVSQLSRAIKREFDPDEPTTLA----IDLMSD 187
| :
Oy 321 P-QATWGCPVNTLRGAGVWFPGPDVTDNPLQSHRLSYVRSRHECKENGHGFMDNKIITTF 379
| :
Db 188 PMNQKGNNPNS-RGVSYVGSDDLARKLDRLQIDLVRHQVVQDGTEFFANRLVLTIF 246
| :
Oy 380 SASNYIAIGSNKGAYIRLNQLMPEHV 406
| :
Db 247 SAPFYCGFDNAAVAAMVYNKLVCSEV 273

RESULT 28
S31088
phosphoprotein phosphatase (EC 3.1.3.16) 1 catalytic chain (clone TOPP4) - Arabidopsis thaliana
M:Alternate names: hypochemical protein Atg39840; protein T517.14
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S31088; T01014; B84822
R:Smith, R.D.; Walker, J.C.
Plant Mol. Biol. 21, 307-316, 1993
A>Title: Expression of multiple type 1 phosphoprotein phosphatases in Arabidopsis thaliana
A:Reference number: S31085; MUID:93144705; PMID:7678768
A:Accession: S31088
A:Molecule type: mRNA
A:Residues: 1-321 <SMS>
A:Cross-references: UNIPROT:P48484; UNIPARC:UPI0000131FB4; EMBL:M93411; NID:g16680; PID submitted to the EMBL Data Library, November 1997
A:Description: Arabidopsis thaliana chromosome II BAC T517 genomic sequence.
A:Reference number: Z14162
A:Accession: T01014
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-321 <ROU>
A:Cross-references: UNIPARC:UPI0000131FB4; EMBL:AC003000; NID:g2642152; PIDN:AAB87136.1; A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Roundley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Umayam, T.; Tallon, L.; Neuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: B84822
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <STO>
A:Cross-references: UNIPARC:UPI0000131FB4; GB:AE002093; NID:g2642169; PIDN:AAB87136.1; G C:Genetics:
A:Gene: Atg39840; T517.14
A:Map position: 2
A:Introns: 73/1: 259/3
C:Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phosphokey words: iron, metalloprotein; phosphoric monoester hydrolase; serine/threonine-specific F:68-136/Domain: phosphoprotein phosphatase homology <PPP>
P:74, 76, 102/Binding site: iron (Asp, His, Asp) #status predicted
F:102, 134, 189, 258/Binding site: zinc (Asp, Asn, His, His) #status predicted
F:105, 135, 282/Active site: Asp, His, Tyr #status predicted
F:106, 231/Binding site: substrate phosphate (Arg) #status predicted

```

Query Match          12.2%; Score 419.5; DB 2; Length 321;
Best Local Similarity 32.3%; Pred. No. 4.4e-19;
Matches 104; Conservative 58; Mismatches 123; Indels 37; Gaps 8;

Oy LVIDPFRKGRGNRIAPHKPVALLIREAAKSLQQL-----PNISPTAVASQQTV 155
   :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db VLDDIRRLTRVRLARPEKQVQLSEA--EIQQLCTYARDIYLQQPNIILEAPIK----I 71
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

Oy CGDLHGKLDDLVLVHLKXGLPSSSNPYVFNDFVDGKRGLEVVLLLSLYLAEPNAVFL 215
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db CGDHHGQSDDLRLFEYVGFPSPSAN-YLFLGDYVDROGSLETICLLAYIKIKPGANFPL 130
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

Oy NRGHEDSVNNAARYGFIREVSCKYPNNHKRIALIDEVRYRLPLGVLSANSVYLIVHGFS 275
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db LRGNHCASINRITIGFDECCRFR--NVRRWVKFTCCENCLPVAAIIDDKILCMHGGLS 187
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

Oy -DSTSLDLIKSIDGKYVSIILRPPLTDEGPIDKTWEOQIPDIMSDPOATGCVPNTLRG 334
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db PDLHLDBDIRLP-----RPTMIDPTGL-----LCDLMSDREKQVKGGMANDRG 232
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

Oy AGWFPGPDVTNPLQRRLSYVIRSHECKENGHEPMHNKITTPSASNYYAIGSNKAY 394
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db VSYTFEGPKVSEFLTAKHDLDVCRAHQVEBDGEVFADRLQIVTFSAPNYCGEFNDAGAM 292
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

Oy IRLNQMLPHVOYISAASOTK 416
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db MSVDENIMCFSFOILKPARKTK 314
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

RESULT 29
T25993
phosphoprotein phosphatase (EC 3.1.3.16) 1 ZK354.9 [similarity] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 05-Oct-2004
C:Accession: T25993
R:Johnson, D.; Wamley, P.; Bradshaw, H.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans coemid ZK354.
A:Reference number: Z20120
A:Accession: T25993
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-310 <JOh>
A:Cross-references: UNIPROT:P91569; UNIPARC:UPI000007FC70; EMBL:U88172; PIDN:AA842261.1
A:Experimental source: Strain Bristol N2; clone ZK354
C:Genetics:
A:Gene: CESP:ZK354.9
A:Map position: 4
A:Introns: 27/2
C:Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phos-
C:Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci-
F:42-110/Domain: phosphoprotein phosphatase homology <PP>
F:48-50, 76/Binding site: iron (Asp, His, Asp) #status predicted
F:76, 108, 158, 250/Binding site: zinc (Asp, Asn, His, His) #status predicted
F:79, 109, 274/Active site: Asp, His, Tyr #status predicted
F:80, 223/Binding site: Substrate phosphate (Arg) #status predicted

Query Match          12.2%; Score 418.5; DB 1; Length 310;
Best Local Similarity 32.2%; Pred. No. 4.9e-19;
Matches 103; Conservative 56; Mismatches 126; Indels 35; Gaps 10;

Oy LPVRKHLDLIDVFRKKGRNLHPKVALLIREAAKSLQQL--NISPTAVASQQTV 155
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db I MPLISRPPIKYSQPAEK-----KFI-----KKFQNLINFMQPMILIRANSPINI 45
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

Oy CGDLHGKLDDLVLVHLKXGLPSSSNPYVFNDFVDGKRGLEVVLLLSLYLAEPNAVFL 215
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db CGDIHGQPSDDLRLFDKNNGPPIHRAN-YLFLGDYVDROGHCHLETILLPLAKVIFPNPFM 104
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

Oy NRGHEDSVNNAARYGFIREVSCKYPNNHKRIALIDEVRYRLPLGVLSANSVYLIVHG-- 273
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

```

Db 105 LRGNHCSLINTKYGFDEGQRRY--NKP8VYVSPFSGVSDMPLTALVGRILCHGVS 162

QY 274 -----FSDSTSLDL-IKSIDRGKYVSIILAPPLTDGEPDLKTEMQOIFDIMSDP-QATMG 326

Db 163 KVVSNFYDVOQINLNTLOLVNSQLRAIKRPFDPNPEPNTLA-----IDLMDSDPTNPGQ 217

QY 327 CVPTLRGAGVFGPVTNTNPLQRHLSVYIRSHCKPNGHPMDNKITITPSASNTYA 386

Db 218 NMPNS-RGVSYYFGSDALKRLDLRQIDLVRAHQVQDGEPPANRLVTLTFSAPYCG 276

QY 387 IGSNKGAYIRLNQNLMPHFV 406

Db 277 QFDNAAVYVYVKNLVCSFV 296

RESULT 30

S32595

phosphoprotein phosphatase (BC 3.1.3.16) 1 catalytic chain - Yeast (*Saccharomyces cerevisiae*)

C:Species: *Saccharomyces cerevisiae*

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C:Accession: S32595; S30854; S50636; E32550

R:Peng, Z.; Wilson, S.B.; Peng, Z.Y.; Schlander, K.K.; Reimann, E.M.; Trumbly, R.J.

J. Biol. Chem. 266, 23796-23801, 1991

A:Title: The yeast Glc7 gene required for glycogen accumulation encodes a type 1 protein

A:Reference number: S32595; M01D:92084672; PMID:160885

A:Accession: S32595

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-312 <FEN>

A:Cross-references: UNIPROT:P32598; UNIPARC:UPI0000052DD3; EMBL:M77175; NID:G172218; PID:R1Mulligan, J.T.; Dietrich, F.S.; Hennessy, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor, submitted to the EMBL Data Library, February 1993

A:Reference number: S30854

A:Accession: S30854

A:Molecule type: DNA

A:Residues: 1-312 <MTL>

A:Cross-references: UNIPARC:UPI0000052DD3; GB:U18916; EMBL:L11120; NID:G1384128; PID:NAR:R1Dietrich, F.S.

A:Description: The sequence of *S. cerevisiae* cosmid 9761, 8196, 9115, 9981, and lambda submitted to the EMBL Data Library, December 1994

A:Reference number: S50636

A:Accession: S50636

A:Molecule type: DNA

A:Residues: 1-312 <DIR>

A:Cross-references: UNIPARC:UPI0000052DD3; EMBL:U18916; NID:G1384128; PIDN:AAC03231.1; R:Ohkura, H.; Kinoshta, N.; Miyatani, S.; Toda, T.; Yanagida, M. Cell 57, 997-1007, 1989

A:Title: The fission yeast dsl2(+) gene required for chromosome disjoining encodes one c

A:Reference number: A32550; M01D:89288305; PMID:2544298

A:Accession: E32550

A:Molecule type: DNA

A:Residues: 1-272, 'V', 274-312 <CHK>

A:Cross-references: UNIPARC:UPI0000175976; GB:M27070

C:Genetics:

A:Gene: SGD:GLC7, DIS281

A:Cross-references: SGD:S0000935; MIPS:YER133W

A:Map position: 5R

A:Introns: 59/3

C:Function:

A:Description: phosphoric monoester hydrolase, serine/threonine-specific phosphatase

C:Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phospho

C:Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci

F:29-288/Domain: phosphoprotein phosphatase homology <PPP>

F:57-125/Domain: phosphoesterase core homology <PPC>

F:63-65, 91/Binding site: iron (Asp, His, Asp) #status predicted

F:91,123,172,242/Binding site: zinc (Asp, Asn, His, His) #status predicted

F:94,124,277/Active site: Asp, His, Tyr #status predicted

F:95,220/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.2%; Score 418.5; DB 2; Length 312;

Best Local Similarity 31.5%; Pred. No. 4, 9e-19;

Matches 100; Conservative 63; Mismatches 125; Indels 29; Gaps 9;

QY 103 NHIDLLIDVFRKKGR--LHPKYVALIREAAK-SLKQLPNISPVSTAVSQQVTCGL 159

Db 10 NIDRLERAGSKPQQQVDLESEIRYLCSKRSIFIQ-----PILBELAPIKICGI 64

QY 160 HGLKDLLVLTKNKLBPSSNPVYNGPDRGKGLVLLLSLYLAFPNVAVLANKN 219

Db 65 HQQYDYLRLPEYGGFPBSN-YLPLADYVVRGKSLTICLLAYKIKYPENPFIILN 123

QY 220 HEDSVNARAGYPIRVSQKYPNHRKILAFIDEVRYMPLGSLNSRVLYVHGFS-DST 278

Db 124 HECASINRYGFIYDECKRY--NTGMKTFD-CNCPRIALIIDBKIFCHHGSLPDLN 180

QY 279 SLDLIKSIDRGKYVSIILAPPLTDGEPDLKTEMQOIFDIMSDPQATMGCVPTLRGAGV 338

Db 181 SWEQIRRVNR-----FTDIPVGLCLDLMSDPDIDVGMSENDGVSFT 225

QY 339 PGPDTNPLQRHLSVYIRSHCKPNGHPMDNKITITPSASNTYAIGSKGAYIRLN 398

Db 226 PGPDVVNRFLQKQDMLCRAHQVDEYEFPSKRLVTLFSAPYCGEFDNAGMMSVD 285

QY 399 NQLMHPVOYISAAGT 415

Db 286 ESLICSF-QILKPAQKS 301

RESULT 31

S24264

phosphoprotein phosphatase (BC 3.1.3.16) 1A catalytic chain - Arabidopsis thaliana

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C:Accession: S24264

R:Ferreira, P.C.G.; Hemerly, A.S.; van Montagu, M.; Inze, D. submitted to the EMBL Data Library, May 1992

A:Description: Molecular characterization of protein phosphatases 1A from *Arabidopsis*.

A:Reference number: S24264

A:Accession: S24264

A:Molecule type: mRNA

A:Residues: 1-312 <FER>

A:Cross-references: UNIPROT:P48482; UNIPARC:UPI000016DBB8; EMBL:Z12163; NID:G16432; PIDN:A:Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phospho

C:Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci

F:64-295/Domain: phosphoprotein phosphatase homology <PPC>

F:64-132/Domain: phosphoesterase core homology <PPC>

F:70,72,98/Binding site: iron (Asp, His, Asp) #status predicted

F:98,130,179,254/Binding site: zinc (Asp, Asn, His, His) #status predicted

F:101,131,278/Active site: Asp, His, Tyr #status predicted

F:102,227/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.2%; Score 418.5; DB 2; Length 312;

Best Local Similarity 30.3%; Pred. No. 4, 9e-19;

Matches 108; Conservative 63; Mismatches 129; Indels 57; Gaps 10;

QY 68 QYQGSARHSVLDHDDVLEBFQDYNALIEPIRKNTDLIDVRRKRG-----NRLH 121

Db 4 QCGQSDPAALDD-----IIR-----LIDYNNPRGRTQAMLNSE 40

QY 122 PKYVALIREAAKSLQKLPNISPVSTAVSQQVTCGLHGLKDLLVLTKNGLBPSSNP 181

Db 41 IRQLCTIVERET---FLQQPNLLEAPRK---ITGDIHQVSDLLRLPRYGGFPPTAN 92

QY 182 YVFNQDPIYDRGKGLVLLLSLYLAFPNVAVLNRGNHEDSVNARAGYPIRVSQKYP 241

Db 93 YLFIDYVDRGQSLTICLLAYKIKYPENPFIIRGNHRCASINRYIGFYDECKRRF-- 150

QY 242 NKRILATIDVYRRLPLGSLNSRVLYVHGFS-DSTSLDLIKSIDRGKYVSIILAPPL 300

Db 151 -SVRLMKVFTFSFNCPLPAAVYIDDKILCMHGAISDLITNVRQIKYK----- 197

QY 301 DQEPDLKTEMQOIFDIMSDPQATMGCVPTLRGAGVFGPVTNTNPLQRHLSVYIRSH 360

Db 198 ---PTDVDSGLICDLILMSDPKSDYKGMGANDRGVSTFGPDKVAFLIKNDMLICRAH 254

QY 361 ECKPENGHEFMNDKITTIFASANTYVAIGSKGAYITLNNQIMPHVQYISAASQTR 417
 DB 255 QVVEDGYEFPADKQVLVTIFSAIPYCGEFDNAGAMMSVDSLMCSF-QILKPADKRRP 310

RESULT 32

T09547
 phosphoprotein phosphatase (EC 3.1.3.16) 1, catalytic gamma chain - alfalfa
 C/Species: Medicago sativa (alfalfa)
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C/Accession: T09547
 R/Vlast: B.; Georadas Toth, B.; Kovacs, I.; Magyar, Z.; Horvath, G.; Bagossi, P.; Gergely
 Arch. Biochem. Biophys. 360, 206-214, 1998
 A/Title: Protein phosphatase 1 catalytic subunit isoforms from alfalfa: Biochemical char
 A/Reference number: Z16730; PMID:99068922; PMID:9851832
 A/Accession: T09547
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1325 <VLS>
 A/Cross-references: UNIPROT:O65845; UNIPARC:UP1000009DAB2; EMBL:AJ002486; NID:G3176073;
 A/Experimental source: subspecies sativa, strain RA3
 C/Genetics:
 A/Genes: PPI gamma
 C/Function:
 A/Description: catalyzes hydrolysis of peptidyl-phosphoserine or -phosphothreonine to re
 C/Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phos
 C/Keywords: phosphoric monoester hydrolase; serine/threonine-specific phosphatase
 F/25-284/Domain: phosphoprotein phosphatase homology <PPP>
 F/53-121/Domain: phosphoesterase core homology <PEC>

Query Match 12.2%; Score 418.5; DB 2; Length 325;
 Best Local Similarity 30.1%; Pred. No. 5.2e-19;

Matches 102; Conservative 66; Mismatches 120; Indels 51; Gaps 9;

QY 90 DIYNAKIELPIRKNIHDLIDVPRKKGRNLRHPKYVALILREAKSLKOLPSS----- 143
 DB 7 DIIRNLIEV-----RNRPG-----KQVUL-----SSEIRQLCNVSRDIFLK 43
 144 -PVSTAVSQVTVCGDLHGKLDLVLVLRNKGLPSSSNPVFNGDPEVDRKGRLEVLTL 202
 DB 44 QPFLDLQLEAPIKCGVHGYSDLLRLFEYGFPPRAN-YLPLGVDVDRKQSLERITCLL 102
 QY 203 LSLIYLPANVFLNRGNHDSVNNARYGPIREYKSPRNKRKILAFIDVYRMLPGSV 262
 DB 103 LAKIKYEPENFLLRGHBCASINRIYGFYDECKRRF--NVLRLMTFTDCSNCLPVAAI 159
 QY 263 LNSRVLLVHGGSF-DSTSLDLKSIDRGKVSILRPLDGEPLDKTEWQIPIDIMWSDP 321
 DB 160 IDKRLICMGGLSPDLHNLQIKNLQ-----PTDVPDTGLCDLLMSDP 204
 QY 322 QAATMGCVPTLRGAGVFGPDVTDNPLQRRHLSYVIRSHCKENGHEFMNDKITTIFSA 381
 DB 205 SKVOGQGMNDKRVSTFGADKITEFLERKIDLDLICRAHQVVDGYEFPANRQVLITTYSA 264
 QY 382 SNTYVAIGSKGAYITLNNQIMPHVQYISAASQTRLSF 420
 DB 265 PNYCGEFDNAGAMMSVDETLMCSF-QILKPADKRRKTLNF 302

RESULT 33

T18936
 phosphoprotein phosphatase (EC 3.1.3.16) 1-beta catalytic chain - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C/Accession: T18936; T25398
 R/McMurray, A.
 submitted to the EMBL Data Library, June 1996
 A/Reference number: Z19047
 A/Accession: T18936
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-401 <WLS>

A/Cross-references: UNIPROT:O18148; UNIPARC:UP100000787E1; EMBL:Z73968; P1DN:CAA98230.1.
 A/Experimental source: clone C05A2
 R/Kelly, P.
 submitted to the EMBL Data Library, June 1996
 A/Reference number: Z20028
 A/Accession: T25398

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-401 <W12>

A/Cross-references: UNIPARC:UP100000787E1; EMBL:Z73977; P1DN:CAA98291.1; GSPDB:GN00023;
 A/Experimental source: clone T28B11
 C/Genetics:

A/Genes: CESP:C05A2.1

A/Map position: 5

A/Insertions: 47/2; 120/1; 159/3; 197/1; 351/3
 C/Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phos

C/Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci

F/87-346/Domain: phosphoprotein phosphatase homology <PPP>
 F/115-183/Domain: phosphoesterase core homology <PEC>

F/121,123,149/Binding site: iron (Asp, His, Asp) #status predicted
 F/149,181,230,305/Binding site: zinc (Asp, Asn, His, His) #status predicted

F/152,230,312/Active site: Asp, His, Tyr #status predicted
 F/153,278/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.0%; Score 413.5; DB 2; Length 401;
 Best Local Similarity 27.7%; Pred. No. 1.4e-18;

Matches 110; Conservative 74; Mismatches 158; Indels 55; Gaps 12;

QY 61 QAAGRKNQOGSAHVSVDKDLVBERFDIYNAKIELPIRKNIHDLIDVPRKKGRNRL 120
 DB 52 QMMRRKNSSSFGELNI-----DQLISTVSVKPMKHTLVTENEL----- 92
 QY 121 HPRYVALILREAKSLKQIPNISPVSTAVSQVTVCGDLHGKLDLVLVLRNKGLPSSSN 180
 DB 93 -RMVCLARQIFMHQPMLELEP-----PLKIGDIGHQPADLRLFNLAGYPPESN 143
 QY 181 PYVNGDPYDRKGRGLEVLTLISLYLAPNNAVFLNRGNHDSVNNARYGPIREYKSKYP 240
 DB 144 -YLFGLDYDRPKSLJETVLLCYKIKYNNFPLRGHBEVANLRIYGFYDECKRRY- 201
 QY 241 RHNKRILAFIDVYRMLPGSVLNSRVLLVHGGSFDS-TSLDLKSIDRGKVSILRPL 299
 DB 202 --SVKLMKCFQDVFNMCPAALIDNKIFCCGGLSNLMSLDQKLSR----- 248
 QY 300 TGBEPLDKTEWQIPIDIMWSDPAT-MGCVPTLRGAGVFGPDVTDNPLQRRHLSYVIR 358
 DB 249 ----PCDVGETGLCOVLSMDPDATVGVNAPR-REVSVFVGVDVLAQGLQKMDLIVVR 303
 QY 359 SHECKENGHEFMNDKITTIFASANTYVAIGSKGAYITLNNQIMPHVQYISAASQ---T 415
 DB 304 GHQVVEDGYEFPGRRLVTVFSAIPYCGEFDNAGAMMSVDSLMCSF-QILKPOSQLVMD 362
 QY 416 KRLSPKQRMGIYESSALKEIAVRMRDRDELDEPKF 452
 DB 363 AALAEKQNNFVANTLGMANIKIKQKP-QPMIDKRRFR 398

RESULT 34

PAEFLA
 phosphoprotein phosphatase (EC 3.1.3.16) 1-alpha-2 catalytic chain - fruit fly (Drosophi
 N/Alternate names: phosphoprotein phosphatase isoform 87b catalytic chain
 C/Species: Drosophila melanogaster
 C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 05-Oct-2004
 C/Accession: S12960; S05524
 R/Dombardt, V.; Axton, J.M.; Barker, H.M.; Cohen, P.T.W.
 FEBS Lett. 275, 39-43, 1990
 A/Title: Protein phosphatase 1 activity in Drosophila mutants with abnormalities in mltc
 A/Reference number: S12960; PMID:91085574; PMID:2175717
 A/Accession: S12960
 A/Molecule type: DNA
 A/Residues: 1-302 <DOM>
 A/Cross-references: UNIPROT:P12962; UNIPARC:UP1000012488F; GB:S55198; GB:S47852; NID:98;
 R/Dombardt, V.; Axton, J.M.; Glover, D.M.; Cohen, P.T.W.

Eur. J. Biochem. 183, 603-610, 1989
 A>Title: Cloning and chromosomal localization of Drosophila cDNA encoding the catalytic
 A:Reference number: S05524; MUID:89377827; PMID:2550221
 A:Accession: S05524
 A:Molecule type: mRNA
 A:Residues: 1-302 <DOM2>
 A:Cross-references: UNIPARC:UPI000012488F; EMBL:X15583; NID:98363; PIDD:CAA33609.1; PID:
 C:Genetic:
 A:Gene: FlyBase: Pp1-87B
 A:Cross-references: FlyBase: FBgn0004103
 A:Map position: 3R 87B-12
 C:Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phos
 C:Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci
 F:28-287/Domain: phosphoprotein phosphatase homology <PPP>
 F:56-124/Domain: phosphoesterase core homology <PEC>
 F:62,64,90/Binding site: iron (Asp, His, Asp) #status predicted
 F:90,122,171,246/Binding site: zinc (Asp, Asn, His, His) #status predicted
 F:93,123,270/Active site: Asp, His, Tyr #status predicted
 F:94,219/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.0%; Score 413; DB 1; Length 302;
 Best Local Similarity 32.3%; Pred. No. 1e-18;
 Matches 108; Conservative 60; Mismatches 128; Indels 38; Gaps 11;

QY 89 GIVNAKIRLPIKKNHIDLLIDVFRKRGKGRNLPKYVALI---LREAAKSLKQLPNISPV 145
 DB 2 GDVNN-----IDSLIRLLEVRGAR-PGRNVQLSGBIRLCLKSRIRFISQPI 49
 QY 146 STAVSQQVTCVGDHKGKLDLVLVLRKNGLPSSSNPYVNGFVDRGKRGLEVLTLTL 205
 DB 50 LLELAPLKLICGDIHQGYDILRLFEYGGFPPESSN-YLFLGIVYDRGKSLFTICLLAY 108
 QY 206 YLAFPAVFLNNGNEDSVNMAKYGIRVSESKYPRNHRKILAFIDVRYRMLPLSGVLS 265
 DB 109 KIKYSNFPLLKRNHCASINRIYGYDCKRKY--SILMKITFD-CFNCIPVAIYDE 165
 QY 266 RVLIVAGFS-DTSLDLTKSIDRGKYVILRPPLDGPPLDKTEWQIIFDIMSDDP-QA 323
 DB 166 KIFCCHGSLPDLTSRQIRLRK-----PTVPPDGLICDLIMSDPDKO 210
 QY 324 TWGCVPTLRGAGVFGPVDYTNFLQHRRLSYVIRSHCKPNGHEPMHNKIITTFPSAN 383
 DB 211 TWGMEEND-RGVFTFGAEVAKFLQKHEFDLICRAHQVVDGYEFAKRLVTLFSAVN 269
 QY 384 YVAGSNKAYIRLNQMLPHFQYISASQTRK 417
 DB 270 YCGEFDNAGAMMSVDDTLWCSP-QILKPADKRRK 302

RESULT 35
 S26225
 phosphoprotein phosphatase (EC 3.1.3.16) 1 - wild cabbage
 N:Alternate names: serine/threonine phosphatase type 1
 C:Species: Brassica oleracea (wild cabbage)
 C:Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
 C:Accession: S26225
 R:Bundle, S.J.; Naarallah, J.B.
 Plant Mol. Biol. 20, 367-375, 1992
 A>Title: Molecular characterization of type 1 serine/threonine phosphatases from Brassic
 A:Reference number: S26225; MUID:93043027; PMID:1330067
 A:Accession: S26225
 A:Molecule type: mRNA
 A:Residues: 1-316 <RUN>
 A:Cross-references: UNIPROT:P48487; UNIPARC:UPI0000131FBB; EMBL:X63558; NID:917881; PIDD:
 C:Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phos
 C:Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci
 F:45-303/Domain: phosphoprotein phosphatase homology <PPP>
 F:73-141/Domain: phosphoesterase core homology <PEC>
 F:79,81,107/Binding site: iron (Asp, His, Asp) #status predicted
 F:107,139,187,262/Binding site: zinc (Asp, Asn, His, His) #status predicted
 F:110,140,286/Active site: Asp, His, Tyr #status predicted
 F:111,235/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.0%; Score 412; DB 2; Length 316;
 Best Local Similarity 30.6%; Pred. No. 1.3e-18;
 Matches 100; Conservative 61; Mismatches 112; Indels 54; Gaps 10;

QY 90 DIVNAKIRLPIKKNHIDLLIDVFRKCR---GNRLHPKVALILREAAKSLKQLPNIS--- 143
 DB 24 DILRLVE-----FRNTRGSGQVH-----LSGG--EIRQLCAVSDI 60
 QY 144 ----PVSTAVSQQVTCVGDHKGKLDLVLVLRKNGLPSSSNPYVNGFVDRGKRGLEVL 199
 DB 61 FLQQLILLELAPIRICGDIHQGYDILRLFEYGGFPPESSN-YLFLGIVYDRGKSLFTI 119
 QY 200 LLLSLVLAFAVFLNNGNEDSVNMAKYGIRVSESKYPRNHRKILAFIDVRYRMLPL 259
 DB 120 CFLVAKIKYPRNFPFLKRNHCASINRIYGYDCKRKY--NRLMKITFD-CFNCIPV 176
 QY 260 GSVLNSRVLIYVAGFS-DTSLDLTKSIDRGKYVILRPPLDGPPLDKTEWQIIFDIMS 318
 DB 177 ALALD-RILCHNGISPLMSLDQIRSLR-----PLDIPDGLVCDLLM 220
 QY 319 SDPQATWCVPTLRGAGVFGPVDYTNFLQHRRLSYVIRSHCKPNGHEPMHNKIIT 378
 DB 221 SDPSGDVAGWGANDRGVFTFGADTVAAEFLQKNDMLICRAHQVVDGYEFAKRLVTV 280
 QY 379 FSAASYVIAIGSNKAYIRLNQMLPHF 405
 DB 281 FSAFNYCEPFDNAGAMMSIDSLWCSP 307

RESULT 36
 PABY3
 phosphoprotein phosphatase (EC 3.1.3.16) PPH3 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein D4421; protein YD8554.08; protein YD8075w
 C:Species: Saccharomyces cerevisiae
 C:Date: 30-Jun-1993 #sequence_revision 10-Feb-1995 #text_change 05-Oct-2004
 C:Accession: S44331; S48762; S55822; S49830; C41525; S67892; S17483
 R:Hoffmann, R.; Jung, S.; Ehrmann, M.; Hofer, H.W.
 Yeast 10, 567-578, 1994
 A>Title: The Saccharomyces cerevisiae gene PPH3 encodes a protein phosphatase with prope
 A:Reference number: S44331; MUID:95028155; PMID:7941742
 A:Accession: S44331
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-308 <HOR>
 A:Cross-references: UNIPROT:P32345; UNIPARC:UPI0000053121
 R:Coester, F.; Jomiaux, J.L.; Goffeau, A.
 submitted to the EMBL Data Library, October 1994
 A:Reference number: S48762
 A:Accession: S48762
 A:Molecule type: DNA
 A:Residues: 1-308 <COS>
 A:Cross-references: UNIPARC:UPI0000053121; EMBL:X82086; NID:9558241; PIDD:CAA57602.1; PI
 R:Coester, F.; Jomiaux, J.L.; Goffeau, A.
 Yeast 11, 673-679, 1995
 A>Title: Analysis of a 32.8 kb segment of yeast chromosome IV reveals 21 open reading fr
 A:Reference number: S55822; MUID:96093910; PMID:7483840
 A:Accession: S55822
 A:Molecule type: DNA
 A:Residues: 1-308 <COM>
 A:Cross-references: UNIPARC:UPI0000053121; EMBL:X82086; NID:9558241; PIDD:CAA57602.1; PI
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
 R:Richards, C.; Harris, D.E.
 submitted to the EMBL Data Library, November 1994
 A:Reference number: S49830
 A:Accession: S49830
 A:Molecule type: DNA
 A:Residues: 1-308 <RIC>
 A:Cross-references: UNIPARC:UPI0000053121; EMBL:Z46796; NID:9577794; PIDD:CAA6797.1; PI
 R:Rome, H.; Carlberg, M.; Hu, G.Z.; Nehlin, J.O.
 Mol. Cell. Biol. 11, 4876-4884, 1991
 A>Title: Protein phosphatase 2A in Saccharomyces cerevisiae: effects on cell growth and
 A:Reference number: S41525; MUID:92017761; PMID:1656215

C:Keywords: glycogen metabolism; heterodimer; iron; metalloprotein; phosphoric monoester
F:25-288/Domain: phosphoprotein phosphatase homology <PP>
F:57-125/Domain: phosphoesterase core homology <PBC>
F:63,65,91/Binding site: iron (Asp, His, Asp) #status predicted
F:91,123,172,247/Binding site: zinc (Asp, Asn, His, His) #status predicted
F:95,124,211/Active site: Asp, His, Tyr #status predicted
F:95,220/Binding site: substrate phosphate (Arg) #status predicted

Query Match 12.0%; Score 411.5; DB 1; Length 333;
Best Local Similarity 31.0%; Pred. No. 1.5e-18;
Matches 100; Conservative 65; Mismatches 133; Indels 25; Gaps 8;

QY 98 LPIRKHHIDLIDYFRKKGRNLHPKYVALI---LKEAASLKLQPLPISVPSAVSQV 154
D 1 MDVEKKNLMDNIIRLLREVRGSK--PGKNVQLTSSBIKLCCKSRERIFLQPILELEAPLK 59
QY 155 VCGDILHGKLDLILVLIHKNLPLPSSNPYVNGPVDVRGKGLFVLLILLILSLYLAFPNAV 214
D 60 ICGDVHQYYDLRLRFYGGFPPEBSN-YLPLGIVDRGKSLSTICLLLYKIKYPENPF 118
QY 215 LNRGNHDSVMNARYGFIREVSXYPNRHRIILAFIDVYRWLPLGSVLSRVLIYVGF 274
D 119 LIRGNHECASINRIYGYDECKRRY-NIKLMKFTFD-CENCLPVAAIIDEKIFCGHGL 175
QY 275 S-DSTSLDLIKSIDRGKYVSLRPLTDSPLDKTEKQQLFDLIMSPPQATMGCVPTLR 333
D 176 SPLQSSMQIRRRIMR-----PTVPDQGLICDLMSPPDQVVGWGENDR 220
QY 334 GAGWFGPDYTDNPLQHRRLSVYIRSHCKPKNGHBFPHNDKIIITPSASNYVYVIGSKGA 393
D 221 GVSTFQPEVYAKFLKHHDDLICRAHQVVEDGIEFFAKQVTLTFSAPNYCGEFDNAGS 280
D 394 YIRLNQLMRPFVQYISASQTK 416
D 281 MMTVDETLMGSP-QILKPADKKK 302

RESULT 58
T19701
phosphoprotein phosphatase (BC 3.1.3.16) C34C12.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 05-Oct-2004
C:Accession: T19701
R:Kerhaw, J.
submitted to the EMBL Data Library, December 1994
A:Reference number: Z19166
A:Accession: T19701
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-382 <WLL>
A:Cross-references: UNIPARC:UPI0000172815; EMBL:Z46996; PIDN:CAA87100.1; GSPDB:GN00021;
A:Experimental source: clone C34C12
C:Genetics:
A:Gene: CBSP:C34C12.3
A:Map position: 3
A:Intons: 7/72; 107/2; 145/1; 179/2; 221/3; 265/3; 292/3; 336/1
C:Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; pho-
C:Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-spec
F:96-356/Domain: phosphoprotein phosphatase homology <PPP>
F:124-192/Domain: phosphoesterase core homology <PBC>
F:130,132,158/Binding site: iron (Asp, His, Asp) #status predicted
F:158,190,240,315/Binding site: zinc (Asp, Asn, His, His) #status predicted
F:161,191,339/Active site: Asp, His, Tyr #status predicted
F:162,288/Binding site: substrate phosphate (Arg) #status predicted

Query Match 11.9%; Score 409.5; DB 1; Length 382;
Best Local Similarity 37.3%; Pred. No. 2.4e-18;
Matches 100; Conservative 45; Mismatches 98; Indels 25; Gaps 7;

QY 141 NISPVSAVSAQVTVCGDLHGKLDLILVLIHKNGLPSSSNPYVNGPVDVRGKGLFVLL 200
D 117 NNPVPS-----PTTICGDIDIGQFYDDL-ELFKYGGVTPNRTKIVFMGDYVDRGHTSLFTV 171

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2006, 19:38:07 ; Search time 163 Seconds
(without alignments)

2861.069 Million cell updates/sec

Title: US-09-463-733-1

Perfect score: 3442

Sequence: 1 MDENAIRAIFIQKMYRRHQ.....VEHDIDPTDCSKVIDPKKS 661

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_05.80.*
1: uniprot_sprot.*
2: uniprot_tramb.*

SUMMARIES

Result No. Score Match Length DB ID Description

| | | | | | | | |
|----|--------|-------|-----|---|--------------|--------|-------------|
| 1 | 3442 | 100.0 | 661 | 1 | RDGC_DROME | P40421 | drosophila |
| 2 | 3121 | 90.7 | 603 | 2 | Q5UID3_DROME | Q5UID3 | drosophila |
| 3 | 1930.5 | 56.1 | 613 | 2 | Q7Q229_ANOGA | Q7Q229 | anopheles g |
| 4 | 1813 | 52.7 | 604 | 2 | Q7QNI0_ANOGA | Q7QNI0 | anopheles g |
| 5 | 1371.5 | 39.8 | 705 | 2 | Q4SOB1_TETNG | Q4SOB1 | tetradon n |
| 6 | 1359 | 39.5 | 757 | 1 | PPE2_MOUSE | PPE2 | mus musculu |
| 7 | 1341 | 39.0 | 707 | 2 | Q61650_CABER | Q61650 | caenorhabdi |
| 8 | 1330 | 38.6 | 753 | 1 | PPE2_HUMAN | PPE2 | homo sapien |
| 9 | 1317 | 38.3 | 707 | 2 | Q01921_CABER | Q01921 | caenorhabdi |
| 10 | 1285.5 | 37.3 | 653 | 2 | Q4R4Y0_MACFA | Q4R4Y0 | macaca fasc |
| 11 | 1271.5 | 36.9 | 653 | 1 | PPE1_HUMAN | PPE1 | homo sapien |
| 12 | 1263 | 36.7 | 683 | 2 | Q9M6F4_FUGRU | Q9M6F4 | fugu rubrip |
| 13 | 1238.5 | 36.0 | 671 | 2 | Q4SHR7_TETNG | Q4SHR7 | tetradon n |
| 14 | 1232 | 35.8 | 797 | 2 | Q4RUI1_TETNG | Q4RUI1 | tetradon n |
| 15 | 1147 | 35.3 | 572 | 2 | Q8MYR2_CABER | Q8MYR2 | caenorhabdi |
| 16 | 839 | 24.4 | 405 | 2 | Q4R441_MACFA | Q4R441 | macaca fasc |
| 17 | 627.5 | 18.2 | 493 | 2 | Q68EP0_XENIR | Q68EP0 | xenopus tro |
| 18 | 618.5 | 18.0 | 897 | 2 | Q4QCW7_XENIR | Q4QCW7 | xenopus tro |
| 19 | 615.5 | 17.9 | 492 | 2 | Q42205_XENIR | Q42205 | xenopus lae |
| 20 | 615.5 | 17.9 | 493 | 2 | Q6GPE6_XENIR | Q6GPE6 | xenopus lae |
| 21 | 614.5 | 17.9 | 474 | 2 | Q7QBI1_ANOGA | Q7QBI1 | anopheles g |
| 22 | 613 | 17.8 | 874 | 2 | Q4RFR2_TETNG | Q4RFR2 | tetradon n |
| 23 | 603.5 | 17.5 | 953 | 2 | Q4Y018_PLACH | Q4Y018 | plasmodium |
| 24 | 594.5 | 17.3 | 923 | 2 | Q81728_TRYIC | Q81728 | trypanosoma |
| 25 | 582.5 | 16.9 | 473 | 2 | Q512P5_ENTHI | Q512P5 | entamoeba h |
| 26 | 581 | 16.9 | 496 | 2 | Q9NEB8_CABER | Q9NEB8 | caenorhabdi |
| 27 | 578 | 16.8 | 829 | 2 | Q4Z563_PLASB | Q4Z563 | plasmodium |
| 28 | 577.5 | 16.8 | 594 | 2 | Q8MOR3_PLAFR | Q8MOR3 | plasmodium |
| 29 | 577.5 | 16.8 | 658 | 2 | Q8IDR7_PLAF7 | Q8IDR7 | plasmodium |
| 30 | 577 | 16.8 | 520 | 2 | Q9YH81_DROME | Q9YH81 | drosophila |
| 31 | 575.5 | 16.7 | 525 | 2 | Q5CJAB_CRYHO | Q5CJAB | cryptospori |

ALIGNMENTS

| RESULT 1 | RDGC_DROME | STANDARD; | PRT; | 661 AA. |
|----------|---|-----------|------|---------|
| AC | P40421; Q9VMA4; | | | |
| DT | 01-FEB-1995 (Rel. 31, Created) | | | |
| DT | 01-FEB-1995 (Rel. 31, Last sequence update) | | | |
| DT | 13-SEP-2005 (Rel. 48, Last annotation update) | | | |
| DE | Serine/threonine protein phosphatase rdgc (EC 3.1.3.16) (Retinal degeneration C protein). | | | |
| GN | Name=rdgc; ORFNames=CG6571; | | | |
| OS | Drosophila melanogaster (Fruit fly). | | | |
| OC | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; | | | |
| OC | Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; | | | |
| OC | Ephydroidea; Drosophilidae; Drosophila. | | | |
| OX | NCBI_TaxID=7227; | | | |
| LN | [1] | | | |
| RP | NCBIOTIDE SEQUENCE (ISORF B). | | | |
| RX | MDLLINE=92263198; PubMed=1316807; DOI=10.1016/0092-8674(92)90230-A; | | | |
| RA | Steele F.R., Washburn T., Rieger R., O'Tousa J.B.; | | | |
| RT | "Drosophila retinal degeneration C (rdgc) encodes a novel | | | |
| RL | serine/threonine protein phosphatase."; | | | |
| RL | Cell 69:669-676(1992). | | | |
| RN | [2] | | | |
| RP | NCBIOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | | | |
| RC | STRAIN=Berkeley; | | | |
| RX | MDLLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185; | | | |
| RA | Adams M.D., Celniker S.B., Holt R.A., Evans C.A., Gocayne J.D., | | | |
| RA | Ananitsides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., | | | |
| RA | George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., | | | |
| RA | Sutton G.G., Wortman J.R., Yandel M.D., Zhang Q., Chen L.X., | | | |
| RA | Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D., | | | |
| RA | Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., | | | |
| RA | Abril J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D., | | | |
| RA | Bellum R.W., Bezu A., Baxendale J., Bayraktaroglu L., Beasley E.M., | | | |
| RA | Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., | | | |
| RA | Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotler P., | | | |
| RA | Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., | | | |
| RA | Cherry J.M., Cline S., Dahlke C., Davenport L.B., Davies P., | | | |
| RA | de Pablo B., Delcher A., Deng Z., Deyan R.A., Dew I., Dietz S.M., | | | |
| RA | Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., | | | |
| RA | Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W., | | | |
| RA | Foster C., Gabrielian A.B., Gary N.S., Gelbart W.M., Glasser K., | | | |
| RA | Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., | | | |
| RA | Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J., | | | |
| RA | Hosht N., Houston K.A., Howland T.J., Wei M.-H., Idegawa C., | | | |
| RA | Jalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., | | | |
| RA | Kamel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z., | | | |
| RA | Laake P., Lai Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X., | | | |
| RA | Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., | | | |
| RA | Merkllov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., | | | |
| RA | Mount S.M., Moy M., Murphy B., Murphy L., Muny D.M., Nelson D.L., | | | |
| RA | Nelson D.R., Nelson K.A., Nixon K., Polard J., Puri V., Reese M.G., | | | |
| RA | Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., | | | |
| RA | Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., | | | |

| | | | | | | | |
|----|-------|------|-----|---|--------------|--------|-------------|
| 32 | 574.5 | 16.7 | 516 | 2 | Q4USE6_THREN | Q4USE6 | theliera a |
| 33 | 573.5 | 16.7 | 534 | 2 | Q96ZNF_PLAFR | Q96ZNF | plasmodium |
| 34 | 570.5 | 16.6 | 483 | 2 | Q5W6J3_ORYSA | Q5W6J3 | oryza sativ |
| 35 | 570.5 | 16.6 | 485 | 2 | Q8H1H4_LYCER | Q8H1H4 | lycopersico |
| 36 | 570 | 16.6 | 484 | 2 | Q9BPW0_HUMAN | Q9BPW0 | homo sapien |
| 37 | 570 | 16.6 | 499 | 1 | PPPS_HUMAN | PPPS | homo sapien |
| 38 | 570 | 16.6 | 499 | 2 | Q53FR0_HUMAN | Q53FR0 | homo sapien |
| 39 | 570 | 16.6 | 499 | 2 | Q53XV2_HUMAN | Q53XV2 | homo sapien |
| 40 | 569 | 16.5 | 479 | 2 | Q64538_RAT | Q64538 | rattus norv |
| 41 | 569 | 16.5 | 499 | 1 | PPPS_MOUSE | PPPS | mus musculu |
| 42 | 569 | 16.5 | 499 | 1 | PPPS_RAT | PPPS | rattus norv |
| 43 | 569 | 16.5 | 499 | 2 | Q68G16_RAT | Q68G16 | rattus norv |
| 44 | 568 | 16.5 | 499 | 2 | Q5R8T2_PONY | Q5R8T2 | pongo pygma |
| 45 | 567.5 | 16.5 | 492 | 2 | Q4PIK6_HUMAN | Q4PIK6 | homo sapien |

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupeki M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RT Science 287:2185-2195(2000).
RN [3]
RP GENOME REANNOTATION, AND ALTERNATIVE SPLICING.
RX MEDLINE:24246069; PubMed:12537572;
RA Miers S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochman S.E.,
Smith C.D., Tudy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bellencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review.";
RT Genome Biol. 3:R85RACH0083.1-R85RACH0083.22(2002).
CC -1- FUNCTION: Phosphatase required to prevent light-induced retinal
degeneration.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
phosphate.
CC -1- COFACTOR: Binds 1 iron ion per subunit (By similarity).
CC -1- COFACTOR: Binds 1 manganese ion per subunit (By similarity).
CC -1- ENZYME REGULATION: Regulated by calcium.
CC -1- ALTERNATIVE PRODUCTS:
CC -1- Event-Alternative splicing; Named isoforms=2;
CC Name=B; Synonym=C;
CC IsoId=P40421-1; Sequence=Displayed;
CC Name=A;
CC IsoId=P40421-2; Sequence=VSP_009324;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Expressed in the visual systems of the fly, as
well as in the mushroom bodies of the central brain.
CC -1- SIMILARITY: Belongs to the PPP phosphatase family.
CC -1- SIMILARITY: Contains 3 EF-hand domains.
CC -1- SIMILARITY: Contains 1 IQ domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
DR EMBL: M69628; AAB00734.1; -; Genomic DNA.
DR EMBL: AE003514; AAF49044.2; -; Genomic DNA.
DR EMBL: AE003514; AA041217.1; -; Genomic DNA.
DR PIR: A42287; A42287.
DR HSSP: Q08209; IAU1.
DR Ensemble: CG6571; *Drosophila melanogaster*.
DR FlyBase: FBgn004366; rdcG.
DR GO: GO:000516; P:calmodulin binding; IDA.
DR GO: GO:0019722; P:calcium-mediated signaling; TAS.
DR GO: GO:0016059; P:deactivation of rhodopsin mediated signaling; IMP.
DR GO: GO:0006470; P:protein amino acid dephosphorylation; IMP.
DR InterPro: IPR011992; EF-hand type.
DR InterPro: IPR002048; EF-hand_Ca_Bd.
DR InterPro: IPR000048; IQ_CaM_Bd_region.
DR InterPro: IPR012008; M-pesterase.
DR InterPro: IPR006186; T_phlasee_apah.
DR Pfam: PF000036; ehand; 3.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00149; Metallophos; 1.
DR PIRSF: PIRSF00912; PPRF; 1.
DR PRINTS: PRO014; STPHPTAS.
DR ProDom: PD000012; EF-hand; 1.

DR ProDom: PD000252; T_phlasee_apah; 1.
DR PROSITE: PS00018; EF_HAND_1; 2.
DR PROSITE: PS50222; EF_HAND_2; 3.
DR PROSITE: PS50096; IQ; 1.
DR PROSITE: PS00125; SER_THR_PHOSPHATASE; 1.
KW Alternative splicing; Calcium; Hydrolase; Iron; Manganese; Manganese;
KW Metal-binding; Protein phosphatase; Repeat; Sensory transduction;
KW Vision.
FT DOMAIN 7 32 IQ.
FT DOMAIN 441 476 EF-hand 1.
FT DOMAIN 526 561 EF-hand 2.
FT DOMAIN 566 601 EF-hand 3.
FT CA_BIND 539 550 1 (potential).
FT CA_BIND 579 590 2 (potential).
FT REGION 105 413 Catalytic.
FT ACT_SITE 220 223 Proton donor (By similarity).
FT METAL 158 158 Iron (By similarity).
FT METAL 160 160 Iron (By similarity).
FT METAL 187 187 Iron and manganese (By similarity).
FT METAL 219 219 Manganese (By similarity).
FT METAL 271 271 Manganese (By similarity).
FT METAL 360 360 Manganese (By similarity).
FT VARSPPLIC 1 58 Missing (in isoform A).
FT /FTid=VSP_009324.
SQ SEQUENCE 661 AA; 75511 MW; A3DC42933B4CCA33 CRC64;
Query Match 100.0%; Score 3442; DB 1; Length 661;
Best Local Similarity 100.0%; Pred. No. 2e-200;
Matches 661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDENAIRAIFQKRYRHOARERMOECNNOIPONTLYASBODAEYKPPNDLIKMP 60
1 MDENAIRAIFQKRYRHOARERMOECNNOIFQNLVYASBODAEYKPPNDLIKMP 60
DB 1 MDENAIRAIFQKRYRHOARERMOECNNOIFQNLVYASBODAEYKPPNDLIKMP 60
QY 61 QAAGRKNQYQSAHYSVLDDKDDLVVEFGDIYNAXIEPIRKNHIDLIDVPRKRGRL 120
61 QAAGRKNQYQSAHYSVLDDKDDLVVEFGDIYNAXIEPIRKNHIDLIDVPRKRGRL 120
DB 61 QAAGRKNQYQSAHYSVLDDKDDLVVEFGDIYNAXIEPIRKNHIDLIDVPRKRGRL 120
QY 121 HPKTYALLIREANGLKQLPNISPVSTAVSQVTCGLHGLDLYLVLRNGLPSSGN 180
121 HPKTYALLIREANGLKQLPNISPVSTAVSQVTCGLHGLDLYLVLRNGLPSSGN 180
DB 121 HPKTYALLIREANGLKQLPNISPVSTAVSQVTCGLHGLDLYLVLRNGLPSSGN 180
QY 121 HPKTYALLIREANGLKQLPNISPVSTAVSQVTCGLHGLDLYLVLRNGLPSSGN 180
121 HPKTYALLIREANGLKQLPNISPVSTAVSQVTCGLHGLDLYLVLRNGLPSSGN 180
DB 121 HPKTYALLIREANGLKQLPNISPVSTAVSQVTCGLHGLDLYLVLRNGLPSSGN 180
QY 181 PVEFGDVPDGRKGLKLVLLLSLYLAFPAVFLNRNHDSSVNNARYGFRVRSYRP 240
181 PVEFGDVPDGRKGLKLVLLLSLYLAFPAVFLNRNHDSSVNNARYGFRVRSYRP 240
DB 181 PVEFGDVPDGRKGLKLVLLLSLYLAFPAVFLNRNHDSSVNNARYGFRVRSYRP 240
QY 191 PVEFGDVPDGRKGLKLVLLLSLYLAFPAVFLNRNHDSSVNNARYGFRVRSYRP 240
191 PVEFGDVPDGRKGLKLVLLLSLYLAFPAVFLNRNHDSSVNNARYGFRVRSYRP 240
DB 191 PVEFGDVPDGRKGLKLVLLLSLYLAFPAVFLNRNHDSSVNNARYGFRVRSYRP 240
QY 241 RNHKKILAFIDVYRWMLPLGSLVNSRVLVVHGSPDSLSLDSIDRGKYSILRPILT 300
241 RNHKKILAFIDVYRWMLPLGSLVNSRVLVVHGSPDSLSLDSIDRGKYSILRPILT 300
DB 241 RNHKKILAFIDVYRWMLPLGSLVNSRVLVVHGSPDSLSLDSIDRGKYSILRPILT 300
QY 241 RNHKKILAFIDVYRWMLPLGSLVNSRVLVVHGSPDSLSLDSIDRGKYSILRPILT 300
241 RNHKKILAFIDVYRWMLPLGSLVNSRVLVVHGSPDSLSLDSIDRGKYSILRPILT 300
DB 241 RNHKKILAFIDVYRWMLPLGSLVNSRVLVVHGSPDSLSLDSIDRGKYSILRPILT 300
QY 301 DGEPLDKTEWQOIFDIMSDDPAATMGCVPTLRGAGVFGPDVTNFIQRRLSYVIRSH 360
301 DGEPLDKTEWQOIFDIMSDDPAATMGCVPTLRGAGVFGPDVTNFIQRRLSYVIRSH 360
DB 301 DGEPLDKTEWQOIFDIMSDDPAATMGCVPTLRGAGVFGPDVTNFIQRRLSYVIRSH 360
QY 361 ECKPVGHEPMNDKTIITPSASNYVAISNSGAVYRLNNOLMPHVQYISASQYRISF 420
361 ECKPVGHEPMNDKTIITPSASNYVAISNSGAVYRLNNOLMPHVQYISASQYRISF 420
DB 361 ECKPVGHEPMNDKTIITPSASNYVAISNSGAVYRLNNOLMPHVQYISASQYRISF 420
QY 421 KQRMVIVSSALKEIAYMRDRDELDEFRYDPKDSGYISISMKCYMENVTLGLP 480
421 KQRMVIVSSALKEIAYMRDRDELDEFRYDPKDSGYISISMKCYMENVTLGLP 480
DB 421 KQRMVIVSSALKEIAYMRDRDELDEFRYDPKDSGYISISMKCYMENVTLGLP 480
QY 481 RLRLDKLAPGDSQKVNTRTLDLDTVILAEADGMSVMDALYANTASVAINITDA 540
481 RLRLDKLAPGDSQKVNTRTLDLDTVILAEADGMSVMDALYANTASVAINITDA 540
DB 481 RLRLDKLAPGDSQKVNTRTLDLDTVILAEADGMSVMDALYANTASVAINITDA 540
QY 541 DNSGRIITDEFTALDILVAHPGAYSKAEMLCKRMMDLNDGVDVNEFLAARLBDL 600
541 DNSGRIITDEFTALDILVAHPGAYSKAEMLCKRMMDLNDGVDVNEFLAARLBDL 600
DB 541 DNSGRIITDEFTALDILVAHPGAYSKAEMLCKRMMDLNDGVDVNEFLAARLBDL 600
QY 601 HRKEQDENIRRSRSTGSPVAKTATPVTLLADKISKNTLVVEHIDPDCESKVIDPPK 660
601 HRKEQDENIRRSRSTGSPVAKTATPVTLLADKISKNTLVVEHIDPDCESKVIDPPK 660
DB 601 HRKEQDENIRRSRSTGSPVAKTATPVTLLADKISKNTLVVEHIDPDCESKVIDPPK 660


```

Qy      661 S 661
Db      661 S 661

RESULT 2
OSUID3_DROME PRELIMINARY; PRT; 603 AA.
ID 05UID3_DROME PRELIMINARY; PRT; 603 AA.
AC 05UID3;
DT 01-FEB-2005 (TEMBLrel. 29, Created)
DT 01-FEB-2005 (TEMBLrel. 29, Last sequence update)
DE 01-FEB-2005 (TEMBLrel. 29, Last annotation update)
DE RH46370p.
GN Name=rdgc;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxId=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkeley;
RA Stapleton M., Carlson J., Chavez C., Friese B., George R., Pacleb J.,
RA Park S., Man K., Yu C., Rubin G.M., Celisner S.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT015959; AAV36844.1; -, mRNA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR004843; M-pesterase.
DR InterPro; IPR006186; T_phthase_apah.
DR Pfam; PF00036; efhand; 3.
DR Pfam; PF00149; Metallophos; 1.
DR PRINTS; PR00114; STPHPTASR.
DR ProDom; PD000012; EF-hand; 1.
DR ProDom; PD000252; T_phthase_apah; 1.
DR SMART; SM00054; Efh; 3.
DR SMART; SM00156; PF2Ac; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_2.
DR PROSITE; PS00125; SER_THR_PHOSPHATASE; UNKNOWN_1.
SQ SEQUENCE 603 AA; 68121 MW; A2FBCA6A32CE2403 CRC64;

Query Match 90.7%; Score 3121; DB 2; Length 603;
Best Local Similarity 99.8%; Pred. No. 5.3e-181;
Matches 602; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      59 MPOAARKNOYQSAHVSVLDDKDLVEERGDIVNAKIELPRKKNHIDLLIVFRKRGK 118
Db      1 MPOAARKNOYQSAHVSVLDDKDLVEERGDIVNAKIELPRKKNHIDLLIVFRKRGK 60

Qy      119 RLHPKTVALLREAAKSLKQLPNISPVSTAVSQOVTVCGLHGLDLDLVTLHKNGLPSS 178
Db      61 RLHPKTVALLREAAKSLKQLPNISPVSTAVSQOVTVCGLHGLDLDLVTLHKNGLPSS 120

Qy      179 SNPYVNGDFVDRGKGLLEVLLLSLYLAFPNAVFLANGNHDVYNAARYGEFIREVSK 238
Db      121 SNPYVNGDFVDRGKGLLEVLLLSLYLAFPNAVFLANGNHDVYNAARYGEFIREVSK 180

Qy      239 YPNHNRKIIAFIDEVYRWLPLGSVLSRSVLIYVGGSDSTSLDLSIDRGKVSILRPP 298
Db      181 YPNHNRKIIAFIDEVYRWLPLGSVLSRSVLIYVGGSDSTSLDLSIDRGKVSILRPP 240

Qy      299 LIDGSEPLDKTEMQOIFDIMMSDPOATMGCVPTLRAGAGVFGPDVTNDFLQRHLSYVR 358
Db      241 LIDGSEPLDKTEMQOIFDIMMSDPOATMGCVPTLRAGAGVFGPDVTNDFLQRHLSYVR 300

Qy      359 SHECKNGHEFMADNKIITIFSSASNTYALGSNKGAVIRLNNQMLPHVOYISAAQTKL 418
Db      301 SHECKNGHEFMADNKIITIFSSASNTYALGSNKGAVIRLNNQMLPHVOYISAAQTKL 360

Qy      419 SPFGKRGIVESSALXELAVMRDRHDELEDEPKYVPOKSGYISISHMCKMKNENTKGL 478
Db      361 SPFGKRGIVESSALXELAVMRDRHDELEDEPKYVPOKSGYISISHMCKMKNENTKGL 420

```

```

Db      361 SPFGKRGIVESSALXELAVMRDRHDELEDEPKYVPOKSGYISISHMCKMKNENTKGL 420
Qy      479 PWRLDKLAPGTDQKVNYNRTLLDLDVILAEADGMSVMDALYANKASIVALENII 538
Db      421 PWRLDKLAPGTDQKVNYNRTLLDLDVILAEADGMSVMDALYANKASIVALENII 480

Qy      539 DADNNGEITLDEPERTIDLLVAMHPGASKAEMLKCRMMNDLNGDKVLEPFLAFLRS 598
Db      481 DADNNGEITLDEPERTIDLLVAMHPGASKAEMLKCRMMNDLNGDKVLEPFLAFLRS 540

Qy      599 DLHREKQDENIRRSSTGRPSVAKTADPVTLLADKISKNTLVESHIDPDCESKVIDP 658
Db      541 DLHREKQDENIRRSSTGRPSVAKTADPVTLLADKISKNTLVESHIDPDCESKVIDP 600

Qy      659 KXS 661
Db      601 KXS 603

RESULT 3
ID 07Q229 ANOGA
ID 07Q229 ANOGA PRELIMINARY; PRT; 613 AA.
AC 07Q229;
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DR ENSANGP000004936 (Fragment).
GN ORFNames=ENSANG00000003837;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OC NCBI_TaxId=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RA The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RA The Anopheles gambiae re-annotation.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC CC
CC EMBL; AAB01008966; EAA12933.2; -, Genomic DNA.
CC GO; GO:0005509; F:calcium ion binding; IEA.
CC GO; GO:0016787; F:hydrolase activity; IEA.
CC CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

Qy      613 AA; 69803 MW; BA0AAEAFEP8D548 CRC64;
Db      601 KXS 603

```

```

Query Match 56.1%; Score 1930.5; DB 2; Length 613;
Best Local Similarity 59.6%; Pred. No. 1e-108;
Matches 366; Conservative 101; Mismatches 130; Indels 17; Gaps 6;

```

```
QY 5 ATRAAIFIOKMYRRHOAREMORRCNMOIFONLEVASBODQAELYKFENDLIKMPQAG 64
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 AMKAAILIQMYRRFLARIEIRRRYWTWTFOSIEVAGBODQRLYNFPAALITHIPETAG 63
QY 65 RKNQVGSAAHVSVLDDKDLVBSFGDIYNAKIELPIRKNHIDLIDVPRKRGKGNLHPKY 124
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 64 RPLDSQNTSRSSSADEGLTGPBERG-YQSPDIFKFLDKKELAVIIDLFRKK-NRLHAKY 121
QY 125 VALILREAKSLKQLEPNISPVSTAVSQOVTVCGLDHGKLDLVLVHNGKLPSSSNPYVF 184
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 VAGILREASTSKRLPNLNQASTAISKQVTCGLDHGKLDLVLVHNGKLPSSSNPYVF 181
QY 185 NGDFVDRGKGLLEVLLLSLYLAPPAVFLNRGNHEDSVNNAKYGFIREVQKTKHNAE 244
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 182 NGDFVDRGKGLLEVLLLSLYLAPPAVFLNRGNHEDSVNNAKYGFIREVQKTKHNAE 241
QY 245 RILAFIDEYRWLPFGSVLNSRVLVHGFSDSTSLDKSIDRGKYSLRPPLTD-- 301
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 242 RLLKLDIVYRWLPFGITVNNRVLVHGGISDSTDLIRSLDRKYSILARPITESTR 301
QY 302 --GEPIDKTEWQOIFDIMSDPOATMGCVPTLRGAGVMPGPDVTNPLQHRRLSYVRS 359
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 302 PGAEIIDKTEWQVFPDILMSDPQHTGCRPNSLRGAGTYFGPDVTSKFLQRYKQLYVRS 361
QY 360 HECKPENGHEPMHNDKIITIFSAISNYAIGSNKGAAYIRLNNQIMPHFVOYISAASQTKLS 419
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 362 HECKPENGHEPMHNDKIITIFSAISNYAIGSNKGAAYIRLNNQIMPHFVOYISAASQTKLS 421
QY 420 FKQRMGIVESSALKELAVRMRDHRLDEPRKYDPKDSGYISISHMCKMENVTGLGP 479
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 422 FQORGLVLESSALRELAAKLRERRIELEREPKSRDPENKGVPLAKMCAIASATSLGP 481
QY 480 WRLLRKLAAPGDS-----QKQVNTLIDLDTVLLEBAQMSVMDALYANKASL 531
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 482 WRLLRKLAAPGDS-----QKQVNTLIDLDTVLLEBAQMSVMDALYANKASL 541
QY 532 VALFIIDADNSGEITLDEFETAILDLVHMPGAYSKAEMLKCRMDLNGDKVDLNF 591
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 542 EALFRLDKDNGSQISLEBFGBACELLRHFP-HNTHBQLDKCRMDINKQGLVDLNF 600
QY 592 LEAFRLSDLHRKEQ 605
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 601 LEFRLCE-NAKEQ 613
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
QYON10 ANOGA PRELIMINARY; PRT; 604 AA.
AC 07ON10;
DT 01-MAR-2004 (T-EMBLrel. 26, Created)
DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
DE ENSANGP0000000059 (Fragment).
GN ORFNames=ENSANG0000000059;
OS Anopheles gambiae str. PRST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
OP (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PRST.
RG The Anopheles gambiae Sequence Committee.
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PRST.
RG The Anopheles gambiae Sequence Committee.
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
   EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
   preliminary data.
```

```
DR EMBL: AA801000995; EAA01910.2; -. Genomic DNA.
DR GO: GO:0005509; F:calcium ion binding; IEA.
DR GO: GO:0016787; F:hydrolyase activity; IEA.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR000048; IQ region.
DR InterPro: IPR004843; M-esterase.
DR InterPro: IPR001861; T.phrase_apha.
DR Pfam: PF00036; ehand; 2.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00149; Metallophos; 1.
DR PRINTS: PR00114; STPHPTASE.
DR ProDom: PD000012; EF-hand; 1.
DR ProDom: PD000252; T.phrase_apha; 1.
DR SMART: SM00054; EFh; 2.
DR SMART: SM00156; PP2Ac; 1.
DR PROSITE: PS00018; EF HAND; UNKNOWN_2.
DR PROSITE: PS00096; IQ; 1.
DR PROSITE: PS00125; SER_THR_PHOSPHATASE; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 604
SQ SEQUENCE 604 AA; 68921 MW; E171D796ADBE450 CRC64;

Query Match
Beet Local Similarity 56.2%; Score 1813; DB 2; Length 604;
Matches 350; Conservative 102; Mismatches 127; Indels 44; Gaps 8;

QY 5 ATRAAIFIOKMYRRHOAREMORRCNMOIFONLEVASBODQAELYKFENDLIKMPQAG 64
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 AMKAAILIQMYRRFLARIEIRRRYWTWTFOSIEVAGBODQRLYNFPAALITHIPETAG 63
QY 65 RKNQVGSAAHVSVLDDKDLVBSFGDIYNAKIELPIRKNHIDLIDVPRKRGKGNLHPKY 124
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 64 RPLDSQNTSRSSSADEGLTGPBERG-YQSPDIFKFLDKKELAVIIDLFRKK-NRLHAKY 121
QY 125 VALILREAKSLKQLEPNISPVSTAVSQOVTVCGLDHGKLDLVLVHNGKLPSSSNPYVF 184
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 VAGILREASTSKRLPNLNQASTAISKQVTCGLDHGKLDLVLVHNGKLPSSSNPYVF 181
QY 185 NGDFVDRGKGLLEVLLLSLYLAPPAVFLNRGNHEDSVNNAKYGFIREVQKTKHNAE 244
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 182 NGDFVDRGKGLLEVLLLSLYLAPPAVFLNRGNHEDSVNNAKYGFIREVQKTKHNAE 241
QY 245 RILAFIDEYRWLPFGSVLNSRVLVHGFSDSTSLDKSIDRGKYSLRPPLTD-- 301
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 242 RLLKLDIVYRWLPFGITVNNRVLVHGGISDSTDLIRSLDRKYSILARPITESTR 301
QY 302 --GEPIDKTEWQOIFDIMSDPOATMGCVPTLRGAGVMPGPDVTNPLQHRRLSYVRS 359
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 302 PGAEIIDKTEWQVFPDILMSDPQHTGCRPNSLRGAGTYFGPDVTSKFLQRYKQLYVRS 361
QY 360 HECKPENGHEPMHNDKIITIFSAISNYAIGSNKGAAYIRLNNQIMPHFVOYISAASQTKLS 419
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 362 HECKPENGHEPMHNDKIITIFSAISNYAIGSNKGAAYIRLNNQIMPHFVOYISAASQTKLS 421
QY 420 FKQRMGIVESSALKELAVRMRDHRLDEPRKYDPKDSGYISISHMCKMENVTGLGP 479
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 422 FQORGLVLESSALRELAAKLRERRIELEREPKSRDPENKGVPLAKMCAIASATSLGP 481
QY 480 WRLLRKLAAPGDS-----QKQVNTLIDLDTVLLEBAQMSVMDALYANKASL 531
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 482 WRLLRKLAAPGDS-----QKQVNTLIDLDTVLLEBAQMSVMDALYANKASL 541
QY 532 VALFIIDADNSGEITLDEFETAILDLVHMPGAYSKAEMLKCRMDLNGDKVDLNF 591
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 542 EALFRLDKDNGSQISLEBFGBACELLRHFP-HNTHBQLDKCRMDINKQGLVDLNF 582
QY 583 DQKVDNLEFLAFLRSLDLHRKEQ 605
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 583 DQKVDNLEFLAFLRSLDLHRKEQ 604
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
QASQBL_TESTING
```

ID Q4SOB1 TETNG PRELIMINARY; PRT; 705 AA.
AC Q4SOB1;
DT 13-SEP-2005 (TREMURel. 31, Created)
DT 13-SEP-2005 (TREMURel. 31, Last sequence update)
DT 13-SEP-2005 (TREMURel. 31, Last annotation update)
DE Chromosome 4 SCAF14533, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00014444001;
OS Tetraodon nigrovittatus (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Mauceli E., Bounneau L., Fischer C., Ozouf-Coestaz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segurens B.,
RA Dastiva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castell V., Katinka M., Vacherie B.,
RA Cnaud C., Duprat S., Brotilier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McSwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mestrov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Lander V., Schachner V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigrovittatus reveals
RT the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL, CAAB01014533; CAP97171.1; -; Genomic_DNA.
FT NON TER 1
FT NON TER 1
SQ SEQIDNCE 705 AA; 81541 MW; B4513BDBA03B1FOA CRC64;
Query Match 39.8%; Score 1371.5; DB 2; Length 705;
Best Local Similarity 39.2%; Pred. No. 1e-74;
Matches 278; Conservative 117; Mismatches 186; Indels 129; Gaps 10;
QY 5 AIRAATFOCKYRHOARRRRCRCMQRQFONLEVASBODQALTYEPFNDLIGMFOAG 64
DB 9 AIRAALLIQRMVQYARLEMRRTYNTITPQSTIEAGBODQIKLYDPFASLMEHFTPAS 68
QY 65 RKN-----OYQGANHYVLDKDKDLVEEFGDVIYAKIELPIRKNHD 106
DB 69 ERLLISHIFPKNEVCCEADWERFYSSVVEPDSYS-----GPRTPFLTICGV 118
QY 107 LLDIVRRKRGKRNLRHYVALLIREAKSLKQLPNISPVSTAVSQOYVTCGLHGLDDI 166
DB 119 KLYEAKFKHQ--OLHARYVQLGERTWRPRLIPNINOVASCONRRTITCGDLHGLEDI 176
QY 167 LVVLRK-----NGLPSSNPYVNGDFDRGKGLVLLLSLYLPNNAVFLNRGN 219
DB 177 LVVYRGLVPAONGPSSSEKPEVNGDFDRGSSSTELLIFGLVLPVNDVHLNRGN 236
QY 220 HEDSVNARVGYRREVSKEYPRNHKILAFIDEVYMLPLGSVINSRVLVHGGFSDSTS 279
DB 237 HEDHIVLRVGYRREVSKEYPRNHKILAFIDEVYMLPLGSVINSRVLVHGGFSDSTS 296
QY 280 LDIKSIDRGKYVILRPP-----TDGEP-----LDKTEMQ 311
DB 297 LDIARVDRHKYVSALRPPKLNQANAKRAGAAGPWEGRRVYTLTYKSTADHSHKRP 356
QY 300 -----TDGEP-----LDKTEMQ 311

DB 357 RSHLSQQLSGGVNMSVEBELKKRRRLAGFDQSYGEPQORSDDSDPESGEATEADHEWK 416
QY 312 QTPDIWSPDQATMGCVPTTAGAGWFGPDVTDNFIQHRRLSYVRSHECKPNGEFPAH 371
DB 417 QVVDLWSPDPMQSGCIPNEVGGGCGYVGPVTEKTLASHDDIQLVRSHECKQEGYEFCH 476
QY 372 DKKIIITFASNYVYAGSNKAGYIRLNQMLPMPVQYISAAGQTKLSKORNGYIYSSA 431
DB 477 DRRVLITFASNYVEVGNRGAIVIRNGHDLIPHPVOY-QASKTCKRLTLKQSGVTERSA 535
QY 432 LKELAVRMDHDEDEBERKXDPKDSGYISISHWKVMEVYKGLPWRLLRDKLAPGT 491
DB 536 LKALKQGLVHNSDLMSAQFQBDPNNVTGMSLRHMASATERVLANGLPVRVLRPQVVSST 595
QY 492 DSGKNTYNTL-DLDTDVILPABADGMSVMDALYANKASLVAFINIIDADNGEITLDE 550
DB 596 ESGMDVYQOWIREFSTTEKLEH-ADN-SILEMTYRNHNLITETPIIIDTHSGLSIPES 653
QY 551 PETAIDLVYAHMPGAYSKAMLEKCRMDNDGKVDLNEPLRAPLSDL 600
DB 654 FHOYTKLSSHLKMEISDKAVTDLAQSIDFNKDSIDINEFMEAFRLVHL 703
RESULT 6
PPE2 MOUSE
ID PPE2 MOUSE STANDARD; PRT; 757 AA.
AC Q35385;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Serine/threonine protein phosphatase with EF-hands-2 (EC 3.1.3.16)
DE (PPEP-2).
GN Name=Ppef2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathu;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RX MEDLINE=97471020; PubMed=9326663; DOI=10.1073/pnas.94.21.11639;
RA Sherman P.M., Sun H., Macke J.P., Williams J., Smallwood P.M.,
RA Nathans J.;
RT "Identification and characterization of a conserved family of protein
RT serine/threonine phosphatases homologous to Drosophila retinal
RT degeneration C";
RL Proc. Natl. Acad. Sci. U.S.A. 94:11639-11644(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6; TISSUE=Retina;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan K.J., Walker J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: May play a role in photoreduction. May

```

CC      dephosphorylate photoactivated rhodopsin. May function as a
CC      calcium sensing regulator of ionic currents, energy production or
CC      synaptic transmission.
CC      -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC      phosphate.
CC      -1- COFACTOR: Binds 1 iron ion per subunit (By similarity).
CC      -1- COFACTOR: Binds 1 manganese ion per subunit (By similarity).
CC      -1- ENZYME REGULATION: Activated by calcium (By similarity).
CC      -1- TISSUE SPECIFICITY: Detected in retina, more specifically in
CC      photoreceptors.
CC      -1- SIMILARITY: Belongs to the PPP phosphatase family.
CC      -1- SIMILARITY: Contains 3 EF-hand domains.
CC      -1- SIMILARITY: Contains 1 IQ domain.
CC      -----
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      -----
CC      EMBL, AF023458; AAB82798.1; -; mRNA.
CC      EMBL, BC027049; AAH27049.1; -; mRNA.
CC      HSSP, P09860; 1DTL.
CC      DR   Ensemble: ENSMUSG0000029410; Mus musculus.
CC      DR   MGI: 1342304; Ppof2.
CC      DR   InterPro: IPR011992; EF-Hand type.
CC      DR   InterPro: IPR002048; EF-hand_Ca_bd.
CC      DR   InterPro: IPR000048; IQ_Cam_Bd_region.
CC      DR   InterPro: IPR004843; M-Pesterase.
CC      DR   InterPro: IPR012008; PPRF.
CC      DR   InterPro: IPR006186; T_phase_apah.
CC      DR   Pfam: PF00036; etand; 3.
CC      DR   Pfam: PF00612; IQ; 1.
CC      DR   Pfam: PF00149; Metallophos; 1.
CC      DR   PIRSF, PIRSF000912; PPRF; 1.
CC      DR   PRINTS, PR00114; STPHPRASE.
CC      DR   ProDom: PD000012; EF-hand; 1.
CC      DR   ProDom: PD000252; T_phase_apah; 1.
CC      DR   SMART, SM00054; EFh; 3.
CC      DR   SMART, SM00015; IQ; 1.
CC      DR   SMART, SM00156; PP2ac; 1.
CC      DR   PROSITE, PS00018; EF_HAND_1; 3.
CC      DR   PROSITE, PS00222; EF_HAND_2; 3.
CC      DR   PROSITE, PS50096; IQ; 1.
CC      DR   PROSITE, PS00125; SSR_THR_PHOSPHATASE; 1.
CC      KW   Calcium; Hydrolyase; Iron; Manganese; Metal-binding;
CC      KW   Protein phosphatase; Repeat; Sensory transduction; Vision.
CC      FT   DOMAIN 21
CC      FT   DOMAIN 572
CC      FT   DOMAIN 656
CC      FT   DOMAIN 696
CC      FT   CA_BIND 669
CC      FT   CA_BIND 709
CC      FT   REGION 128
CC      FT   ACT_SITE 241
CC      FT   METAL 179
CC      FT   METAL 181
CC      FT   METAL 208
CC      FT   METAL 240
CC      FT   METAL 292
CC      FT   METAL 492
CC      SEQUENCE 757 AA; 86645 MW; 821B3D1061AC0C8 CRC64;
CC
CC      Query Match      39.5%; Score 1359; DB 1; Length 757;
CC      Blast Local Similarity 38.9%; Pred. No. 6,5e-74;
CC      Matches 278; Conservative 113; Mismatches 200; Indels 124; Gaps 7;
CC
CC      3 ENAIRPAIFIGYRRHQARRRMRRCNMWQIFONLEVASBODQAEIYFNDLIKHMQA 62
CC      :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
CC      17 EKAFPAFALIQWYRYVYARLEMRRCNWNIFQSIYAGQDQVYGLHFFSYVDHFP 76
CC      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC      63 AGKKVQYGSAAHVSVLTDKNDLIVERFGI-----VNAKIELPIKNIHDLIVER 113
CC      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

| | | | |
|--|---|--|-------------------------------|
| Db | 77 | SHHEEDFLNRKPTTERFADVTEBGGDFESITVPDSTYGPLSLPFLDLPHATALVEAR | 136 |
| Qy | 114 | KKKGRLHPKVVALLIREAAKSLKQLPNISPVSTAVSQQVTVCGDLHGKLLDLVLVLRN | 173 |
| Db | 137 | LRQ--QLHARVYANLVEYTRFKLADLPNINRSTCVSEBVTYCGDLHGQDLDDLIPIFYRN | 194 |
| Qy | 174 | GLPSSSNVYVNGDVVDGKRGLEVLLLSLYLAPNAVFLNRGNHDSVMNARYGFTK | 233 |
| Db | 195 | GLPSEPERAYVFNQDVGDKGSVEVLTMYLFAFMLYYPKSFHLNRGNHDEHVLNLRGYFTK | 254 |
| Qy | 234 | FVESYCPANHKRIILAFIDEVYRWMLPLGSVLRNRYIVHGQFSDSTSLDIKSIDREKYS | 293 |
| Db | 255 | EVMKRYIKHGKILLTLDQVFCMLPLATLVDEKVLVHGQVSDKTDLSLTLAKIDRHKIVS | 314 |
| Qy | 294 | ILR----- | 296 |
| Db | 315 | TMRCTKRSENRBEQKQKNDQTSQGQPTWFLPQSRLEPSSPFLHSGFAYKAGRC | 374 |
| Qy | 297 | -----PLPLT-----GEP | 304 |
| Db | 375 | SIPGSPNSKELSRGQVRYSVDLEBQCRQAGFLGIREKESLPLADACVADGGCV | 434 |
| Qy | 305 | LDKT--EMQOILFDIMSPDPOATMGCVPTLRGAGVFGQVDYDNFLQHRRLSVYIRSHC | 362 |
| Db | 435 | LEPPIEEKQVVDILMSDPAALGGKANAVRGGGCTFGQDVTRLEMKKILLQULLRSHCC | 494 |
| Qy | 363 | KPNGHFPMNDKILITTFASNVYYAIGSKGAYIRLNQMLPMFVQYISAASTKRLSFQ | 422 |
| Db | 495 | KPEGEGFCHNRKVLITFASNVYYEGSKRGAVYVKGPAITPIIVQ--QANKATHLITMQ | 553 |
| Qy | 423 | RMGIVSSALXELAVRMRDHDELEDFRKYDPKDGYISISHWCKMENVTGLPML | 482 |
| Db | 554 | RISRVEBALRALRQKLPASHSDLLVEFRKRPDSSGVITLSDMTAVAVSLHGLPRM | 613 |
| Qy | 483 | LKDLCAPCTDQKVVYNNRTLLDLDVYLBAADQMSWDALYANKASVALFNITDAQN | 542 |
| Db | 614 | LRPOLVNSSANVVEYRWLBSLAKEQ--LSRNNIGSSILKELKYRNRSNLETFRIIDSDH | 672 |
| Qy | 543 | SGEITLDFEFALDILVAHMPGAYKAEMLTEKRMMDLNGDKVDLNEPLFAFR | 597 |
| Db | 673 | GGFISLDFRQTWKLPSSHMSIDITDDGICDLARSIDPRKQGHIDINEFLBAFRL | 727 |
| RESULT 7 | | | |
| 0616S0 CAEBR PRELIMINARY; PRT; 707 AA. | | | |
| ID | 0616S0 | CAEBR PRELIMINARY; | PRT; 707 AA. |
| AC | 0616S0 | | |
| DT | 25-OCT-2004 | (TReMBLrel. 28, Created) | |
| DT | 25-OCT-2004 | (TReMBLrel. 28, Last sequence update) | |
| DT | 25-OCT-2004 | (TReMBLrel. 28, Last annotation update) | |
| DE | Hypothetical protein CBG15166. | | |
| GN | Name=CBG15166; | | |
| OS | Caenorhabditis briggsae. | | |
| OC | Bukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea; | | |
| OC | Rhabditidae; Pelodertinae; Caenorhabditis. | | |
| OX | NCBI_TaxId=6238; | | |
| RN | [1] | | |
| RP | NUCLEOTIDE SEQUENCE. | | |
| RG | The C.briggsae Sequencing Consortium; | | |
| RL | Submitted (SEP-2003) to the EMBL/Genbank/DBJ databases. | | |
| CC | -!- CAUTION: The sequence shown here is derived from an | | |
| CC | EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is | | |
| CC | preliminary data. | | |
| DR | EMBL; | CAACO100069; | CAB69136.1; -, Genomic DNA. |
| DR | GO; | GO:0005509; | F-actin monomer binding; IEA. |
| DR | GO; | GO:0016787; | F-actin monomer binding; IEA. |
| DR | InterPro; | IPR002048; | EF-hand. |
| DR | InterPro; | IPR000048; | IQ region. |
| DR | InterPro; | IPR004843; | M-dehydrogenase. |
| DR | InterPro; | IPR012008; | p53. |
| DR | InterPro; | IPR006186; | T-actin monomer binding; IEA. |
| DR | PIfam; | PF00036; | ethanol dehydrogenase; IEA. |
| DR | PIfam; | PF00612; | IQ; 1. |

DR Pfam; PF00149; Metallophos; 1.
 DR PIRSF; PIRSF000912; PPERF; 1.
 DR PRINTS; PRO00114; STEPHNTSR.
 DR ProDom; PD000012; EF-hand; 1.
 DR ProDom; PD000252; T_phthase_apah; 1.
 DR SMART; SM00054; EPh; 2.
 DR SMART; SM00156; P2Ac; 1.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN_2.
 DR PROSITE; PS00125; SER_THR_PHOSPHATASE; UNKNOWN_1.
 DR Hypothetical protein.
 DR SEQUENCE 707 AA; 79840 MW; 4ADE2CA1D186C630 CRC64;
 Query Match 39.0%; Score 1341; DB 2; Length 707;
 Best Local Similarity 42.4%; Pred. No. 7.3e-73;
 Matches 269; Conservative 118; Mismatches 192; Indels 56; Gaps 12;
 QY 6 IRAAIFIQKWRHQAARREMOQRNNOIFONTLYASEBODALYKFFNDLIKMPQAAG 65
 DB 83 IKSALIQKWRRCARLEARRATQIFALEYAEQDQKLYDFADVIRAMABENGK 142
 QY 66 KNOYQG--SAHVSVDL--KDLVSEFGDIVNAKIR-----LPINK 102
 DB 143 TGGDNERNSPMLASHYAKPSLMDSESGEYVKMLDTSFGNVVDNRYKGPVLSPLDK 202
 QY 103 NHIDLDIVPRKRGKRLHPKYVALILREAAKSLKQLPNIPSVSTAVSQQVTCGDLHGK 162
 DB 203 POYAKMTEAF--KMKVLIHPKVIKTLMLHEARKIFKSMPSVSRISTISNOITTCGDLHGK 260
 QY 163 LDDLVLVHKNGLPSSNNPVPNGDVPDGRKGLBYLLLSLYLAFPNAVFLRNGHED 222
 DB 261 PDDLCLILKNGVPVDNPIYFNGDVPDGRGGSIVLCVLFALVIDPMSTYLRNHNH 320
 QY 223 SVNNARYGTRIEVSKYTPNNHRIILAFIDETVYMLPLGVSANRVLIVHGSEDSSTSLD 282
 DB 321 H1MNLRYGPIKELSTYKDLSTPTITRLIRVFSWMLIATIVDKDIFVVGGISDQTEVER 380
 QY 283 IKSIDGKVSITLRPLTDG-----BPLDKTEWQQLFDIWNSSPOATMCGVPTLHG 334
 DB 381 LDKIPHRQSVLRKPVNKMDSAEKENVQSYDEMKQMDIMWSPKNGCWPVVFPG 440
 QY 335 AGVWPGPDYDNLQHRILSYVIRSHCKPNCHGFHNDKIITIFASNYAIGSNKAY 394
 DB 441 GGSYFADITASTLEHGFRLLVIRSHCKEGEGEFSHNTCTLVFASNYTEGSRGAY 500
 QY 395 IRLNQLMHPVOYISAAQTRKLS-FKRMGIIVESSALKEALVRNRDRHLEDEFRKY 453
 DB 501 VKFYGKSRHFPQYM--ASKTRKSTLRRLRGLVSESAVLEKRLKLSFTPLTQKEFENV 558
 QY 454 DPRDSGYISISHCKVMENVTKGLPWRLLRDLAIGTSSQK-VATNRLLDLDTVYIL 512
 DB 559 DLKSGNLPILKMSCEVERITGLNLPWIALAPVATLSDGKRVNKKDRKIAQVGTHA 618
 QY 513 AEAAGWSVDALYANKASLVAIFNIIIDANGSEITLDEFEITALLVAHMPGAYSK---- 568
 DB 619 QEKD---LYESLYRHKSTLETFLFRMDKNSGVSKKEFLDACEVL-----GKTYRPLQ 670
 QY 569 ----AEMLEKCRMDLNGDKVDLNLPLRAFLSD 599
 DB 671 TDYIAQIAES---IDFNKQGFIDLNLLEAFRLVD 702
 RESULT 8
 ID PPER2_HUMAN STANDARD; PRT; 753 AA.
 AC 014830; 014831;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 13-BFP-2005 (Rel. 48, Last annotation update)
 DE Serine/threonine protein phosphatase with EF-hands-2 (EC 3.1.3.16)
 DR (PPERF-2).
 OS Name=PPERF2;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
 OC Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE (ISOFORMS PPERF-2(S) AND PPERF-2(L)).
 RC TISSUE=Retina;
 RX MEDLINE=97471020; PubMed=9326663; DOI=10.1073/pnas.94.21.11639;
 RA Sherman P.M., Sun H., Macke J.P., Williams J., Smallwood P.M.,
 RA Nathans J.;
 RT Identification and characterization of a conserved family of protein
 RT serine/threonine phosphatases homologous to Drosophila retinal
 RT degeneration C.";
 RT Proc. Natl. Acad. Sci. U.S.A. 94:11639-11644(1997).
 CC -1-FUNCTION: May play a role in phototransduction. May
 CC dephosphorylate photoactivated rhodopsin. May function as a
 CC calcium sensing regulator of ionic currents, energy production or
 CC synaptic transmission.
 CC -1-CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1-COPFACTOR: Binds 1 iron ion per subunit (By similarity).
 CC -1-COPFACTOR: Binds 1 manganese ion per subunit (By similarity).
 CC -1-ENZYM REGLATION: Activated by calcium (By similarity).
 CC -1-SUBCELLULAR LOCATION: Localized to photoreceptors, PPERF-2(L) is at
 CC least 2 fold more abundant in rod inner segments than in the outer
 CC segments.
 CC -1-ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named Isoforms=2;
 CC Name=PPERF-2(L);
 CC IsoId=014830-1; Sequence=Displayed;
 CC Name=PPERF-2(S);
 CC IsoId=014830-2; Sequence=VSP_005103, VSP_005104;
 CC -1-TISSUE SPECIFICITY: Retinal specific (Probable).
 CC -1-SIMILARITY: Belongs to the PPP phosphatase family.
 CC -1-SIMILARITY: Contains 3 EF-hand domains.
 CC -1-SIMILARITY: Contains 1 IQ domain.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AF023456; AAB82796.1; -; mRNA.
 CC EMBL; AF023457; AAB82797.1; -; mRNA.
 CC HSSP; P09860; 1DTL.
 CC Ensembl; ENSG00000156194; Homo sapiens.
 CC HGNC; HGNC:9244; PPERF2.
 CC MIM: 602255; -
 CC GO; GO:0004722; P:protein serine/threonine phosphatase activity; TAS.
 CC GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
 CC InterPro; IPR011992; EF-Hand type.
 CC InterPro; IPR002048; EF-hand Ca bd.
 CC InterPro; IPR000048; IQ Cam bd region.
 CC InterPro; IPR004843; M-Pesterase.
 CC InterPro; IPR012008; PPERF.
 CC InterPro; IPR006186; T_phthase_apah.
 CC Pfam; PF00036; ehand; 3.
 CC Pfam; PF00612; IQ; 1.
 CC Pfam; PF00149; Metallophos; 1.
 CC PIRSF; PIRSF000912; PPERF; 1.
 CC PRINTS; PRO00114; STEPHNTSR.
 CC ProDom; PD000012; EF-hand; 1.
 CC ProDom; PD000252; T_phthase_apah; 1.
 CC ProDom; PD00054; EPh; 3.
 CC SMART; SM00015; IQ; 1.
 CC SMART; SM00156; P2Ac; 1.
 CC PROSITE; PS00018; EF_HAND_1; 2.
 CC PROSITE; PS00222; EF_HAND_2; 3.
 CC PROSITE; PS00096; IQ; FALS_NG.
 CC PROSITE; PS00125; SER_THR_PHOSPHATASE; 1.
 CC Alternative splicing; Calcium; Hydroxylase; Iron; Manganese;
 CC Metal-binding; Polymorphism; Protein phosphatase; Repeat;
 CC Sensory transduction; Vision.

```

FT DOMAIN 21 46 IQ.
FT DOMAIN 568 603 EF-hand 1.
FT DOMAIN 652 687 EF-hand 2.
FT DOMAIN 692 727 EF-hand 3.
FT CA_BIND 665 676 1 (Potential).
FT CA_BIND 705 716 2 (Potential).
FT REGION 128 540 Catalytic.
FT ACT_SITE 241 241 Proton donor (By similarity).
FT METAL 179 179 Iron (By similarity).
FT METAL 181 181 Iron (By similarity).
FT METAL 208 181 Iron and manganese (By similarity).
FT METAL 240 240 Manganese (By similarity).
FT METAL 292 292 Manganese (By similarity).
FT METAL 488 488 Manganese (By similarity).
FT VARSPPLIC 588 598 IRTSDMAAVE -> ASSQLCTYQOK (in isoform
PEPF-2(S)).
FT VARSPPLIC 599 753 /FTID-VSP_005103.
FT VARSPPLIC 120 120 Missing (in isoform PEPF-2(S)).
FT VARIANT 120 120 /FTID-VSP_005104.
FT S -> R.
SQ SEQUENCE 753 AA, 86431 MW, 289D4AC240193A0D CRC64,
/FTID-VAR_010230.

Query Match 38.6%; Score 1330; DB 1; Length 753;
Best Local Similarity 38.1%; Pred. No. 3,7e-72;
Matches 278; Conservative 109; Mismatches 185; Indels 158; Gaps 11;

3 ENAIRAATITQKMYRRHQAARERQRCNQIQTONTLEYASBODQAEIYKFFNDLIKMPQA 62
17 ERAFKAALIORMYRRYVARLEEMRRRCTSIFQSISYAGQOQVVKLHDFSYLMOH----- 72
63 AGRKXQYQSAHVSVLDDKDDLYEERGD-----IYNAKIEIP 99
73 -----FIRSSH---NDRDFLRITTEBFRPADDSMKKCSDEYSIEVDSYTGSELSP 122
100 IRKNMIDLLIDYFRKKRGRNLHPKYVALIRKAAKSLKQLPNISPYTAVSQOVCYCDL 159
123 LIPDHATALVEAFRLMQ--QLHARYLNLLEYFKKLVDLPINRSTCYSEIITVCGDL 180
160 HGKLDLVLVHKNGLPSSSNPYVNGDVYDRGKGLVLLLSLYLAFNPAVFLNRGN 219
181 HGQLDLIFIFYKNGGLPSPERSVVFNGDVRGKGSVEILMLIFAPMLVYPKPFHLNRGN 240
220 HEDSVNARYGFRREVESKYPRNHRLLAFIDSVYRWLGLSVTLNRVLYVGGSPDS 279
241 HSDHWNALYGFLEKVMNKKYVHGKELKTLQVFCMLPLATLIDKVLILHGVSDDTD 300
280 LPLIKSIDRGKYVSLR-----PPV- 299
301 LELDLIERSKIYSTMRCKTROKSEKOMEKRRANKQSSAOGPIPWFLPESRSLPESPLR 360
300 -----TQGEPLDK----- 307
361 LGSYKAQKTRSSSSIPCSGLDRELSRQVRSVLELERCQOAGLTVGEKEPSSRA 420
308 -----TEMQOIPDIWSDPQATMGCPNPLLRAGVAFPGDYDYNPLQRR 352
421 SRADSEAGELRKPTQESMRQVVDILMSDPAQSGCAANTIRGGCYFGSDVYQQLQKXN 480
353 LSYVIRSHCKPQNGEFPMNDNKITITFSASNYAATRNKGAAYRLNQLMPFVQYISA 412
481 LQFLIRSHCKPQNGEYFCNKRKVLITFSASNYEVSNGAYVKGALPALTPHVOY-QAN 539
413 SOTKRLSPKQMGIVSSSAIKELAVWRDHRDELDEFRKYDPKQSGYISHHCKVMEN 472
540 KYTHLTLMQRISLRVESALRALREKLFAHSSDLSEFFKHADKGLITLSDMAAAVRS 599
473 VTKLGLPMTLLRDKLAPGTDSCVNTNRLDLDLTVIIIEABDQGSVNDALYANKASLY 532
600 VHLGLPMTLLRDKLAPGTDSCVNTNRLDLDLTVIIIEABDQGSVNDALYANKASLY 532
533 ALFNIDIDANSGETLIDDEFETALIDLVAHMPGAYSAEMLEKC-----RWMDLNGDGKXD 587

```

```

DB 659 TIFRIIDSHSGFSLDSERQWTWLFSSHM-----NIDITDDCICDLARSIDFNKGCHID 713
QY 588 INEFLEAPRL 597
DB 714 INEFLEAPRL 723

RESULT 9
001921_CABEL
ID 001921_CABEL PRELIMINARY; PRT; 707 AA.
AC 001921; 017047;
DT 01-JUL-1997 (TRMBLrel. 04, Created)
DT 01-OCT-2001 (TRMBLrel. 18, Last sequence update)
DT 13-SEP-2005 (TRMBLrel. 31, Last annotation update)
DE Phosphatase with ef hands protein 1, isoform a (Protein phosphatase
DE with EF-hands).
GN Name=pef-1; ORFNames=F23H11.8;
OS Caenorhabditis elegans.
OC Rhabditidae; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;
OC Rhabditidae; Pelodetidae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN (1)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Br1etol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT Investigating biology."
RL Science 282:2012-2018 (1998).
(2)
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97471020; PubMed=9326663; DOI=10.1073/pnas.94.21.11639;
RA Sherman P.M., Sun H., Macke J.P., Williams J., Smallwood P.M.,
RA Nathans J.;
RT "Identification and characterization of a conserved family of protein
RT serine/threonine phosphatases homologous to Drosophila retinal
RL degeneration C."
RL Proc. Natl. Acad. Sci. U.S.A. 94:11639-11644 (1997).
DR EMBL, AF023389; AAC7113.2; -; Genomic_DNA.
DR EMBL, AF023454; AAB82794.1; -; mRNA.
DR PIR, T34072; T34072.
DR PIR, T42239; T42239.
DR HSSP, P36873; 1176.
DR Ensembl, F23H11.8; Caenorhabditis elegans.
DR Wormbase, MBgene0003969; pef-1.
DR WormPep; F23H11.8a; CE27939.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR011992; EF-Hand type.
DR InterPro; IPR002048; EF-hand_Ca_bd.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR004843; M-pesterase.
DR InterPro; IPR012008; PEPF.
DR InterPro; IPR006186; T_phase_apah.
DR Pfam; PF00036; ehand; 2.
DR Pfam; PF00612; IQ_1.
DR Pfam; PF00149; Metallophos; 1.
DR PIRSF; PIRSF000912; PPEF; 1.
DR PRINTS; PR00114; STEPHATASE.
DR ProDom; PD000012; EF-hand; 1.
DR ProDom; PD000252; T_phase_apah; 1.
DR SMART; SM00054; EFh; 3.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00156; PZAC; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_2.
DR PROSITE; PS00125; SER_THR_PHOSPHATASE; UNKNOWN_1.
KV Complete proteome.
SQ SEQUENCE 707 AA, 80330 MW, 1943CE21BD0A7EF7 CRC64;

Query Match 38.3%; Score 1317; DB 2; Length 707;
Best Local Similarity 42.1%; Pred. No. 2,1e-71;
Matches 266; Conservative 116; Mismatches 200; Indels 50; Gaps 12;

```


RT "Identification and characterization of a conserved family of protein
RT serine/threonine phosphatases homologous to Drosophila retinal
RT degeneration C.";
RT Proc. Natl. Acad. Sci. U.S.A. 94:11639-11644(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.
RC TISSUE=Fetal brain;
RC MEDLINE=97358589; PubMed=9215685; DOI=10.1093/hmg/6.7.1137;
RA Montini E., Rugat E.I., van de Vosse E., Andolfi G., Mariani M.,
RA Puca A.A., Conales G.G., den Dunnen J.T., Balibio A., Franco B.,
RT "A novel human serine-threonine phosphatase related to the Drosophila
RT retinal degeneration C (rdgc) gene is selectively expressed in sensory
RT neurons of neural crest origin."
RL Hum. Mol. Genet. 6:1137-1145(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=retina;
RC MEDLINE=98104127; PubMed=9430683; DOI=10.1074/jbc.273.3.1462;
RA Huang X., Honkaniemi R.E.;
RT "Molecular cloning, expression, and characterization of a novel human
RT serine/threonine protein phosphatase, PP7, that is homologous to
RT Drosophila retinal degeneration C gene product (rdgc)."
RL J. Biol. Chem. 273:1462-1468(1998).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC PubMed=15772651; DOI=10.1038/nature03440;
RA Rose M.T., Graham D.V., Coffey A.J., Scherer S., McElay K., Muzny D.,
RA Platerer M., Howell G.R., Butrows C., Bird C.P., Frankish A.,
RA Lovell F.L., Howe K.L., Ashurst J.L., Fulton R.S., Sudbrak R., Wen G.,
RA Jones M.C., Hultine M.E., Andrews T.D., Scott C.E., Sealie S.,
RA Ramser J., Whitaker A., Deadman R., Carter N.P., Hunt S.E., Chen R.,
RA Cree A., Gunaratne P., Havlak P., Hodgson A., Metzger M.L.,
RA Richards S., Scott G., Steffen D., Sodergren E., Wheeler D.A.,
RA Worley K.C., Alencough R., Ambrose K.D., Anselmi L., Axtell S.,
RA Ashwell R.I., Babage A.R., Baggsley C.L., Ballal A.A., Banerjee R.,
RA Barker G.E., Barlow K.F., Barrett I.P., Bates K.N., Beare D.M.,
RA Beasley H., Beasley O., Beck A., Bethel G., Blechschmidt K., Brady N.,
RA Bray-Allen S., Bridgman A.M., Brown A.J., Brown M.J., Bonin D.,
RA Bruford E.A., Buhay C., Butch P., Burford D., Burgess J., Buttrill W.,
RA Burton J., Bye J.M., Carder C., Carrel L., Chako J., Chapman J.C.,
RA Chavez D., Chen E., Chen G., Chen Y., Chen Z., Chinault C.,
RA Ciccolicola A., Clark S.Y., Clarke G., Clee C.M., Clegg S.,
RA Clerc-Blankenburg K., Clifford K., Cobley V., Cole C.G., Conquer J.S.,
RA Corby N., Connor R.B., David R., Davies J., Davis C., Davis J.,
RA Delgado O., Dehaene D., Dharm P., Ding Y., Dinh H., Dodsworth S.,
RA Drepper H., Dugan-Rocha S., Dunn A., Dunn M., Durbin K.J., Dutta I.,
RA Faulkner L., Francis R., Frankland J., Fraser A.E., Galoczky P.,
RA Gilbert J., Gill R., Gloeckner G., Gregory S.G., Gribble S.,
RA Griffiths C., Grocock R., Gu Y., Gwilliam R., Hamilton C., Hart E.A.,
RA Hawes A., Heath P.D., Heilmann K., Hennig S., Hernandez J.,
RA Hinemann B., Ho S., Hoffe M., Howden P.J., Huckle R.J., Hume J.,
RA Hunt P.J., Hunt A.R., Isherwood J., Jacob L., Johnson D., Jones S.,
RA de Jong P.J., Joseph S.S., Keenan S., Kelly S., Kershaw J.K., Khan Z.,
RA Kioschke P., Klages S., Knights A.J., Koshira A., Kover-Smith C.,
RA Laird G.K., Langford C., Lawlor S., Leverish M., Lewis L., Liu W.,
RA Lloyd C., Lloyd D.M., Lonsdale H., Loveland J.E., Lovell J.D.,
RA Lozowal J., McLaren S., McMurray A., Meldal P., Meltlinger T.,
RA Milne S., Miner G., Mistry S.L., Morgan M., Morris S., Mueller I.,
RA Mullikin J.C., Nguyen N., Nordesiek G., Nyakatura G., O'dell C.N.,
RA Okunou G., Palmer S., Pandian R., Parker D., Parrish J.,
RA Pasternak S., Patel D., Pearce A.V., Pearson D.M., Pelan S.B.,
RA Perez L., Porter K.M., Ramsey Y., Reichwald K., Rhodes S.,
RA Ridler K.A., Schlessinger D., Schuler M.G., Sehra H.K.,
RA Shaw-Smith C., Shen H., Sheridan R.M., Showkhan R., Stuce C.D.,
RA Smith M.L., Sohrbreck D., Taber P.E., Taudien S., Taylor I.,
RA Teague B., Thomas K., Thorpe R., Timms K., Tracey A., Trevanion S.,
RA Tromans A.C., d'Urso M., Verdusco D., Villaseña D., Waldron L.,
RA Wall M., Wang Q., Warren G., Warruzo G.L., Wei X., West A.,
RA Whitehead S.L., Whiteley M.N., Wilkinson J.E., Willey D.L.,
RA Williams G., Williams L., Williamson A., Williamson H., Wilming L.,
RA Woodmansey R.L., Wray P.W., Yen J., Zhang J., Zhou J., Zoghbi H.,
RA Zorilla S., Buck D., Reinhardt R., Pousetta A., Roenthal A.,
RA Lehnach H., Weindl A., Minx P.J., Hillier L.W., Willard H.F.,
RA Wilson R.K., Waterston R.H., Rice C.M., Vaudin M., Coulson A.,
RA Nelson D.L., Weinstein G., Sulston J.E., Durbin R., Hubbard T.,
RA Gibbs R.A., Beck S., Rogers J., Bentley D.R.;
RT "The DNA sequence of the human X chromosome."
RL Nature 434:325-337(2005).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC TISSUE=Testis;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.P., Jordan B., Moore T., Max S.I., Wang J., Hsieh P.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Tashyuk S., Carlini P., Prange C.,
RA Rahn S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmitt J., Myers R.W.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalski U., Smalusi D.E.,
RA Scherch A., Schin J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP SPLICING ISOFORM(S) THAT ARE POTENTIAL NMD TARGET(S).
RC PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;
RA Hillman R.T., Green R.B., Brenner S.R.;
RT "An unappreciated role for RNA surveillance."
RL Genome Biol. 5:RESEARCH008.1-RESEARCH008.16(2004).
CC -1- FUNCTION: May have a role in the recovery or adaptation response
CC of photoreceptors. May have a role in development.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- COFACTOR: Binds 1 iron ion per subunit (By similarity).
CC -1- COFACTOR: Binds 1 manganese ion per subunit (By similarity).
CC -1- ENZYME REGULATION: Activated by calcium.
CC -1- BIOPHYSICOCHEMICAL PROPERTIES:
CC pH dependence:
CC Optimum pH is 8.0;
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Name=1;
CC IsoId=O14829-1; Sequence=Displayed;
CC Name=1A;
CC IsoId=O14829-2; Sequence=VSP_005098;
CC Name=1B;
CC IsoId=O14829-3; Sequence=VSP_005099;
CC Name=2;
CC IsoId=O14829-4; Sequence=VSP_005100; VSP_005101;
CC Note=May be produced at very low levels due to a premature stop
CC codon in the mRNA, leading to nonsense-mediated mRNA decay;
CC Name=3;
CC IsoId=O14829-5; Sequence=VSP_005102;
CC Note=May have no functional significance;
CC -1- TISSUE SPECIFICITY: Detected in retina and retinal derived Y-79
CC retinoblastoma cells. Also found in fetal brain.
CC -1- SIMILARITY: Belongs to the PPP phosphatase family.
CC -1- SIMILARITY: Contains 3 EF-hand domains.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

use as long as its content is in no way modified and this statement is not removed.

CC EMBL; AF023455; AAB82795.1; -; mRNA.
CC EMBL; X97867; CAAG6461.1; -; mRNA.
DR EMBL; AF027977; AAC05825.1; -; mRNA.
DR EMBL; AL096700; CA142857.1; -; Genomic_DNA.
DR EMBL; Z94056; CA142857.1; JOINED; Genomic_DNA.
DR EMBL; Z94056; CA142777.1; -; Genomic_DNA.
DR EMBL; AL096700; CA142777.1; JOINED; Genomic_DNA.
DR HSSP; O82040; 1K9U.
DR Ensemble; ENSG00000086717; Homo sapiens.
DR HGNC; HGNC:9243; PEPF1.
DR MIM; 300109; -.
DR GO; GO:0004722; F:protein serine/threonine phosphatase activity; TAS.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR InterPro; IPR011992; EF_Hand type.
DR InterPro; IPR002048; EF_Hand_Ca_bd.
DR InterPro; IPR000048; IQ_Cam_Fd_region.
DR InterPro; IPR004843; M-pesterase.
DR InterPro; IPR006186; T_phatase_apah.
DR Pfam; PF00036; ehand; 3.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00149; Metallophos; 1.
DR PRINTS; PIRSF000912; PEPF; 1.
DR PRINTS; PR00114; STPHPTASE.
DR ProDom; PD000012; EF_hand; 1.
DR ProDom; PD000252; T_phatase_apah; 1.
DR SMART; SM00054; Eph; 3.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00156; PP2Ac; 1.
DR PROSITE; PS00018; EF_HAND_1; 2.

Query Match 36.9%; Score 1271.5; DB 1; Length 653;
Best Local Similarity 40.3%; Pred. No. 1.1e-68;
Matches 238; Conservative 129; Mismatches 204; Indels 49; Gaps 11;

CC 3 ENAIRAIFLOKWRHRRHQRREMRQRCNMQIFONLEAYASBODQALYKFPNDLIKMPQA 62
DB 14 DTSLRALILIQNYRGKAKLAKRQHVALTIFOSIEYADQGMQSTFFSFMLENYTHI 73
QY 63 AGRKNQYQSAAHVSLDDKD--DLVR--EFGDIVNA-KIELPIRKNHIDLLIDVFRKRG 117
DB 74 HKEELERINQSLSESDQMDRMDYVDISIDVDSYNGRPLQPLCTDIDLLLEAFKQO- 132
QY 118 NRHPKYVALLLEBAKSLKQLNISPVSVAVSQVTVGCGDKLDDLLVLAHKGLPS 177
DB 133 -IIHAHYVLEVLKETKVKLQKMPNFTHIQSPSEKVTICGDLHGKLDLFLIYKXGLPS 191
QY 178 SSNPVYFNGDFVDRGKRGLEVLILLSLYLAFPNNAVFLNKGHEDSVNARVGFIREVES 237
DB 192 ERNPVYFNGDFVDRGKRSIEIIMILCSPLYVYNDLHNRGNHEDFMNLRVGFTEIELH 251
QY 238 KYPRNKRILAFIDEVYRMPLGSSVLSRVLLVHGGSFSDSLDKSIDRGKYSILRP 297
DB 252 KYTLHGKRILQILEBEFYAMLPIGTIVDNELIVHGHSFTDNLHLHVRNRNKKSLIP 311
QY 298 PL-----TDGEP---LDKTEHQOIFDIMSDDQATMGCV 328
DB 312 PIETNNDHDTDSKKNKGVTFNNAHRLKTNSTPETHLTERHMQIIDILMSDRGKNGCF 371
QY 329 PNTLRGAGVWFGPDVTNDFLQRRHLSVIRSHCKPKNGHFMMDNKITITPSASNYAIG 388
DB 372 PNTCRGGCCYFGSDVYSKILANKQLKMLHSHCKKEGIEICDQKAVITIPSSANYEBS 431
QY 389 SNKGAYIRLNNQMLPHEVQYISAASQTKLSF--KQRMGIVSSALKELAVEMRHRDE 445
DB 432 SNKGAYIKLCSGTTPRPFQY----QVTKATCFQLRQVRVTMNSAKILIREVISRKS 487
QY 446 LDEBFKRYDPKDSGYISISHWCKVMENVTLGLPRLRLRKALPAGDSDQVYNNRTLDL 505

DB 488 LTPAPQLQDHRKSGKLSVSGMAFCMENIIGLNLPMRSLSNVLNIDQNGVNEYMSSFQNI 547
QY 506 DTLD-VILEABADGMSVMDALYANKASLVAILFNIIDNSGEITLDFETAIDLLVAHMPG 564
DB 548 RLEKPYQVQEAHS--TLVETLYRSDLEITFNADIDTGHSLISVEBFAMWKLFPSSHYNV 604
QY 565 AYSKAEMLKRCNMWDNGDGKVDLNFLEAFRLSDHRE 604
DB 605 HIDDQVNRKLANIMIDNKDGSIDFNEFLKAFYV--VHRVY 642

RESULT 12
09W6R4_FUGRU
ID 09W6R4_FUGRU PRELIMINARY; PRT; 683 AA.
AC 09W6R4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE Protein phosphatase 1.
GN Name=PPF-1;
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxId=31033;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99263230; PubMed=10330123;
RA Brunner B., Todd T., Lenzner S., Stout K., Schulz U., Rogers H.-H.,
RA Kalscheuer V.M.;
RT "Genomic structure and comparative analysis of nine Fugu genes:
RT conservation of synteny with human chromosome Xp22.2-p22.1.";
RL Genome Res. 9:437-448(1999).
DR EMBL; AF146687; AAD28796.1; -; Genomic_DNA.
DR HSSP; P02589; ITNX.
DR Ensemble; STNFRUG00000152266; Fugu rubripes.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0016787; F:calcium ion binding; IEA.
DR InterPro; IPR002048; EF_Hand.
DR InterPro; IPR011992; EF_Hand type.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR004843; M-pesterase.
DR InterPro; IPR012008; PEPF.
DR InterPro; IPR012008; PEPF.
DR Pfam; PF00036; ehand; 3.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00149; Metallophos; 1.
DR PRINTS; PIRSF000912; PEPF; 1.
DR PRINTS; PR00114; STPHPTASE.
DR ProDom; PD000012; EF_hand; 1.
DR ProDom; PD000252; T_phatase_apah; 1.
DR SMART; SM00054; Eph; 3.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00156; PP2Ac; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_2.
DR PROSITE; PS00096; IQ; 1.
DR PROSITE; PS00125; SBR_THR_PHOSPHATASE; UNKNOWN_1.
SQ SEQUENCE 683 AA; 78631 MW; EDEB07661E30892 CRC64;

Query Match 36.7%; Score 1263; DB 2; Length 683;
Best Local Similarity 38.7%; Pred. No. 3.8e-68;
Matches 260; Conservative 126; Mismatches 218; Indels 68; Gaps 9;

QY 2 DENAIRAIFLOKWRHRRHQRREMRQRCNMQIFONLEAYASBODQALYKFPNDLIKMPQA 61
DB 14 EKTALKAALVLIQWRWRYARLREMRRTYWNIFOSIEYAGDQQLSSFFSFMLENYTHI 73
QY 62 AAGRK----NOYQSAHVSLDDKD--DLVEEFGDIVNAKIELPIRKNHIDLLIDVFRKK 115
DB 74 LKNGGPDLSKLMDSASDPWLENENCYNLVTVPEYTGRLRSLPPLSVSDMNLGAFKQ 133
QY 116 RGNRLHPKTVALLLEBAKSLKQLNISPVSVAVSQVTVG--GDHLGKLDLLVLAHKNG 174

```

DB      134 Q--TLFARVYLQILHETRKILKOMPNIILHSTTYTKEITICGADLLGQIDDLILYKNG 191
      175 LPSSNPVYFNGDVPDRGKRGLEVLILLSLYAFPNVAFLNKGNHDSVMNARYGPIRE 234
      192 LPSEAFVYFNGDVPDRGKNSVEMVTLFAYLLLYPDYVHNLKNGNHDHMLNRYGFTKE 251
      235 VESKYPANRKLIAFLDEYVYRMPLASVANSRLVYHGFPSDSTSLDLKSIDRGKYVI 294
      252 VMOKYIHGCEILQLPQDVFSLLPIATIIIDGKLLVHGGISDQTDLPFLSLERHKIKSA 311
      295 LRPPLTDE----- 303
      312 LRPQVTVFQSLGSSQVYKIRSTDSSRPSRSSHNIQRTPHGSGVQLPFLDSLBA 371
      304 ---PLDKTEMOQFDIMMSDPOATMGCVPTNLGAGVFGPDVTDNLFQRHLVYIRSH 360
      372 PSPQHOEWEKQIVDILMSDPKTLKCCPTFGGCGYFGPDVTRLLQHGLQLIRSH 431
      361 ECKPNHGFPHMDKIIITTSASNYVAIGSKAYIILNNOVMPHYQYISASQTKLGF 420
      432 ECKQBYELCHSQVITTSASNYBEGSNRGAYIVGRLMRFYQY-QVSRSTKLT 490
      421 KQPMGIVSSALKEIAVRMDHDEDEDEFRKYDPDQSGYISISHCKYMENTYKLGPM 480
      491 TQVRVRAEGSAFPALEKLFTHRSSEILGFGQYDHNNTGIIIVSAQVLETGLRLDPM 550
      481 RLIRDKLAPGTSQKYNVNRFTLDLDTVYLEAADMVMDALYANKASLVAIFNIIDA 540
      551 RLIRPPLVRLASGVRDYQSCFEDMEPGV-PQVEVT-PNLADLFRYRKDIETIFSIIDK 608
      541 DNGSEITLDEFTAIIDLVAHMPGAYSKAEMLEKCMMDLNGGKDLAEFLAFLSLD 600
      609 DQGLISIEFRTWHLFSAHLGIKIDNRAIDDLARSIDFNKGIDFNEFLAFAV--V 666
      601 HRKQODENIR 612
      667 HLDNDQPVNR 678

QY      601 HRKQODENIR 612
DB      667 HLDNDQPVNR 678

RESULT 13
Q4SHR7_TETNG PRELIMINARY; PRT; 671 AA.
ID Q4SHR7_TETNG PRELIMINARY; PRT; 671 AA.
AC Q4SHR7_TETNG PRELIMINARY; PRT; 671 AA.
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DE Chromosome 5 SCAPI4581, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00018033001;
OS Tetradon nigrovittide (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteloscromi;
OC Actinopterygii; Neopterygii; Teleostei; Buteloscromi;
OC Acanthopterygii; Acanthopterygii; Percomorpha; Tetradontiformes;
OC Tetradontidae; Tetradontidae; Tetradon.
OC NCBI_TaxID=99883;
RX [1] YANKASLVAIFNIIDANSGEITLDEFTAIIDLVAHMPGAYSKAEMLEKCMMDLNGG 584
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Coataz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segurens B.,
RA Dasilva C., Sainoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castell L.V., Katinka M., Vacherie B.,
RA Biemont C., Skali Z., Catolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellie M., Volf J.N., Guigo R., Zody M.C., Mestrov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laude V., Schacher V., Queller F., Saurin M., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RA "Genome duplication in the teleost fish Tetradon nigrovittide reveals
RT the early vertebrate proco-karyotype.";
RL Nature 431:946-957(2004).

```

```

RN      [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAB01014581; CAP9815.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 671
SQ SEQUENCE 671 AA; 77077 MW; 54AE1BF22697E7D CRC64;

Query Match      36.0%; Score 1238.5; DB 2; Length 671;
Best Local Similarity 37.7%; Pred. No. 11e-66;
Matches 254; Conservative 122; Mismatches 212; Indels 85; Gaps 7;

QY      5 AIRAIFIOKRYRQAREREMORCNMOIFQNLBYASRQDQALYKFFNDLKHMPQAG 64
      1 AIIKAIVLQRMFRKIMAHLEMRRTYVNIPOSIEYASRQDQALQSFSPMLNITHLNG 60
      65 RF---NOYQSAHVSVIDKD-DLYBFGDIYAKIELPIKXNIDLLIDVPRKGN 118
      61 SGPDLISKLMDSASDSMMNSNCYFVSPEYTGPRLSPLSLPDMNALLSAFQRQ-- 118
      119 RLHPKYVALIIRBAKSLKQLPNISPVSTAVSQVTCGDLHGKLDLIVLHNGLPBS 178
      119 TLHARYVQLLETREKLLKQMPNIVHLSITYYKRIITGDLHGLDILLIYKNGLPBA 178
      179 SNPYFNGDVPDRGKRGLEVLILLSLYAFPNVAFLNKGNHDSVMNARYGPIREVESK 238
      179 EFPYFNGDVPDRGKNSIEVVLLFAYLLLYPDYVHNLKNGNHDHMLNRYGFTSEVMOK 238
      239 YPRNHRKLIAFLDEYVYRMPLASVANSRLVYHGFPSDSTSLDLKSIDRGK----- 290
      239 YKHGCEILQLPQDVFSLLPVATIIDGKLLVHGGISDQTDLPFLSALERHVSSEKLR 298
      291 -----YSLRPL----- 299
      299 HFQIPINRKRCKLRQDSTANSSSSSSSSSSSSISLCKTPCHTPRHPPSSPTLS 358
      300 -----TDGSPDKTEMOQFDIMMSDPOATMGCVPTNLGAGVFGPDVT 344
      359 CGSAPQLPFLDSLAPSPQHOEWEKQIVDILMSDPKTLKCCPTFGGCGYFGPDVT 418
      345 DNFQRHLVYIRSHCKPNHGFPHMDKIIITTSASNYVAIGSKAYIILNNOVMPHY 404
      419 RLLAQGQLIRHSHCKQEGYEICHGQVITTSASNYBEGSNRGAYIVGRLMRFYQY 478
      405 FYQYISASQTKLGFQKPMGIVSSALKEIAVRMDHDEDEFRKYDPDQSGYIS 464
      479 FYQY-QVSRFTKLTQVRVRAEGSALRALEKLFTHRSSEILGFGQYDHNNTGIIIVS 537
      465 HVCYMENTYKLGPMRLRLDCLAQSDQKYNVNRFTLDLDTVYLEAADMVMDAL 524
      538 EMAQVLETGLRLDPMRLRLRLAIDLAGRYHSCFEDMEPGVPIVQVTNLA--SL 595
      525 YANKASLVAIFNIIDANSGEITLDEFTAIIDLVAHMPGAYSKAEMLEKCMMDLNGG 584
      536 FYRRKDIETIFSIIDKDSGLISIEFRTWHLFSAHLGKIDNRAIDDLARSIDFNKG 655
      585 KYDLNFELEAPRL 597
      656 NIDFNEFLAFAV 668

QY      585 KYDLNFELEAPRL 597
DB      656 NIDFNEFLAFAV 668

RESULT 14
Q4RUI1_TETNG PRELIMINARY; PRT; 797 AA.
ID Q4RUI1_TETNG PRELIMINARY; PRT; 797 AA.
AC Q4RUI1_TETNG PRELIMINARY; PRT; 797 AA.
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Chromosome 12 SCAPI4996, whole genome shotgun sequence.

```

(fragment).
 ORFNames=GSTENG0028998001;
 GN Tetradodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthopterygii; Acanthopterygii; Percormorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetradodon.
 NCBI_TaxId=99883;
 [1]
 NUCLEOTIDE SEQUENCE.
 RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
 RA Mauceli B., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Micaud S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segurens B.,
 RA Basilia C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castelli V., Katicina M., Vacherie B.,
 RA Blemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
 RA Cnaud C., Duprat S., Brocletier P., Coutanceau J.P., Gouy J.,
 RA Parra G., Lardier G., Chapple C., McKernan K.J., McGowan P., Bosak S.,
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Meitrov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Lander V., Schachter V., Queller F., Saurin W., Scarpelli C.,
 RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,
 RT "Genome duplication in the teleost fish Tetradodon nigroviridis reveals
 RT the early vertebrate proto-karyotype.";
 RL Nature 431:946-957(2004).
 RN [2]

NCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead Institute Centre for Genome Research;
 RC Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 RL -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; CA001014996; CAG08121.1; -; Genomic_DNA.
 FT NON TER 1 797
 FT 1 797
 SQ SEQUENCE 797 AA; 91973 MW; 89F01F058C820730 CRC64;

Query Match 35.8%; Score 1232; DB 2; Length 797;

Best Local Similarity 32.5%; Pred. No. 3.5e-66;

Matches 261; Conservative 135; Mismatches 190; Indels 218; Gaps 13;

5 AIRAATFIOKMYRHOARERORCMQVQFONLEVASSEDOAEVLYRPFNDLIGKMPQAG 64
 1 ALPAATVIOQMYQYARSKMRKCTWHIQTSTIYSEQAOIKLSNPLNTLMNTPSSN 60
 65 RKNQYGSATVSLDD--KDDLYEEFQDLYNAKI-----ELPIKKNHIDLIDVPR 113
 61 ERLULI---SHIFRENEVCRAEMERYFCYKNIEVPETYSGPHLTPFLTVGQAVGVAFR 117
 114 KRGKGNLHPKTYALLIREAKSKIKQLPNISPVSTAVSQOYTV-----156
 118 NKR--QLHSRYIIQLLETWKLRLMPLNIRISTCHSKETITCGELQEGFGNGLCREIV 175
 157 -----GDHGLKLDLIVLHK-----172
 176 GFTLPKTVLTWVAAPIGDHLGQJEDLLLFYVRNPLPLTFMFSPTVFLAVTYBOTSN 215
 173 -----NGLPSSNPVYFNGDVFVGRKGLSEVLLLSLYLAFNAVFLNR 217
 236 NNNKNPRVFNKALQNGMPLEKPYFNGDFVGRGDSIIIVILVPLAVYSQDYIYLR 255
 218 GNHEDSVNARYPFIREVSKYPRNKRRIIAFIDEVYRMLPLGSLVNSRVLVHGGPSDS 277
 296 GNHEDIIIVNRYPTREKLVKTHQGRKRIILQKIFSWPLATVLDQKVLVHGGISDL 355
 278 TSLDLKSIDRGKYVSLRPP-----LITGSEPL 305
 356 TDLGLAKLDRNHYVSLRPPKRGYVSACTISDSVDDEVCSKIFQRTSLTYARPL 415
 306 D-----K 307
 416 GRTTSFQNRSLQDFSDRIRLPAENELRRRRRPFIFHLSSEKTIINPAASDVKSBSLK 475

308 TEMQOIFDMMSPQATMGCVPTLIRGAGVWFGPDVTIDPRLQORRLSYVIRSHCKPNGH 367
 476 DEKQILDLIMSDPMIQDCIFIRVWGGCTWGPDTTEERLANKMIQLIIRSHCKQBG 535
 368 EFMHDKII-----TFPSANYYAIGSKGAYIRLNQI 401
 536 EFCNKRKVIQGLPGPGPRVLPGHQLKTVPLQVLLTFSSNYYDVGSNRGAVVVLGPNL 595
 402 MPRFVYISAAGTKRLSKFRKMGIVSSALRELAVRMGRDHEDEBRKDPKDSGII 461
 596 VPRVIOY--QANSWTRELSMROSIGRTERSALKVIREQLPAHKSDDLISAKQFPTENTGV 654
 462 SISHWKWMENTVKGELPRLIRDKLAPGTDSQ--KYNVRRITDILNPTVILREADGMSV 520
 655 SLCDMAVAESVWHELCPWRMLSSQSTTKSGDMSTYHWPDEL--AIKGNIDGLQ 711
 521 -----MDALYANKASIVAFNIIDANGSEITLDEPFIADLLVAMPAGVSKAEMLEKC 575
 712 TLVRLHETLYRHRSTLETIFRIIDKNSGFIEMEDRKTKLISLYLKNRISDALSDLA 771
 576 RMDLNGDKVLDNFELEAFRLSD 599
 772 VTIDSQDSIDIDEFMAFRILD 795

RESULT 15

ID 08MYR2 CAEBL PRELIMINARY; PRT; 572 AA.

AC 08MYR2;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DT Phosphatase with ei hande protein 1, isoform b.

GN Name=pef-1; ORFNames=F23H11.8;

GN Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

NCBI_TaxId=62339;

[1] NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=Bristol N2;

RX MEDLINE=99069613; PubMed=9851916;

RG The C. elegans sequencing consortium;

RT "Genome sequence of the nematode C. elegans: a platform for

investigating biology."

RL Science 282:2012-2018 (1998).

DR EMBL; AF003389; AAM22065.1; -; Genomic_DNA.

DR HSSP; P36873; 1IT6.

DR Ensemble; F23H11.8; Caenorhabditis elegans.

DR WormBase; WBGene0003969; pef-1.

DR WormPeP; F23H11.8b; CR30662.

DR GO; GO:0005509; F:calcium ion binding; IEA.

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR InterPro; IPR011992; EF-Hand type.

DR InterPro; IPR002048; EF_Hand_Ca_bd.

DR InterPro; IPR004843; M-pesterase.

DR Pfam; PF00036; ehand; 2.

DR Pfam; PF00149; Metallophos; 1.

DR PRINTS; PR00114; STRPHRTASE.

DR SMART; SM00054; EFh; 3.

DR SMART; SM00156; PP2AC; 1.

DR PROSITE; PS00018; EF_HAND; UNKNOWN 2.

DR PROSITE; PS00125; SER_THR_PHOSPHATASE; UNKNOWN 1.

KM Complete proteome.

SO SEQUENCE 572 AA; 65150 MW; 504039A0685C98CD CRC64;

Query Match 33.3%; Score 1147; DB 2; Length 572;

Best Local Similarity 41.0%; Pred. No. 3.3e-61;

Matches 227; Conservative 106; Mismatches 174; Indels 46; Gaps 10;

81 KDDLYEEFQDLYNAKI-----LPIRKKNHIDLIDVPRKKGNNRLH 121

```

Db      27 KPSLMSBGEIVYKMLBETSPTNVVIDIDRNVYKGPPTLSLPLDKPQVAKMEAF--KVNKVLH 84
Qy      122 PKYVALLIIEBAKSKLKQLPNISPVSTAVSQOVTTCGDLHGKLDLLVYTHKGLPSSNP 181
Db      85 PKYVALLIIEBAKSKLKQLPNISPVSTAVSQOVTTCGDLHGKLDLLVYTHKGLPSSNP 144
Qy      182 YFNGDFVDRGKGLVLLLSLYLAFPAVAFNAGNHEDSVNNAARYGFIREVSCKYR 241
Db      145 YIFNGDFVDRGKGLVLLLSLYLAFPAVAFNAGNHEDSVNNAARYGFIREVSCKYR 204
Qy      242 NHRKILAFIDEVRYMLPLGSVTVNSRLVYHGFSDSTSLDLSIDRGKYSILRPPLTD 301
Db      205 LSTPTLRLLDEVFSWPIPIATIIDRIDFVYHGGISDQTEVSKLDKIPRHRFQSVLAPPVNK 264
Qy      302 GEPBLOK-----TEWQOIFDIMMSDPQATMGCVPTTLGAGVWRPDTDTDFLQHRHS 354
Db      265 GMSBSEKENSANVNDKQMLDIMMSDPKQKQKQWPFVFGGSGYFGADITLAFLEHGRR 324
Qy      355 VYIRSHCEKPNGBEFMDNKITIFGASNYVAIGSNKAYIR-LNNQLMHPFOYISAAS 413
Db      325 LVIRSHCEKPNGBEFMDNKITIFGASNYVAIGSNKAYIR-LNNQLMHPFOYISAAS 382
Qy      414 QTKRLS-FKQRMGIVSSALKELAVMRDHRDELDEFRKYPDKOSGYISISHCKWMEN 472
Db      383 KTHRSKSTLERLGVVESAVKELKELSGFHTDLQKFEIEMDIKRSGLPLKMSDQYER 442
Qy      473 VTKGLPMLLKDKLAP-GTDSQKVNRYNTLDDLTVDVLEKADQMSWMDLYANKSL 531
Db      443 ITGLNLPMTALAPKQATLSBDGCVYKEDRRILAQVGTHAQEKD---IVESILYRKSTL 499
Qy      532 VAIFNIIDNDSGEITLDFETPIDLLVAHMPGAYSKAEM-----LEKRMMDLNGDGKY 586
Db      500 ETILFRMDKDNQSGVSMKEPIADACEVL-----GKTKRPLQTDYISQIAESIDPNQGT 554
Qy      587 DINEFLAEFRSLD 599
Db      555 DINEFLAEFRSLD 567

RESULT 16
Q4R441_MACFA PRELIMINARY; PRT; 405 AA.
ID Q4R441_MACFA PRELIMINARY; PRT; 405 AA.
AC Q4R441;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Testis cDNA clone: QcSA-12561, similar to human protein phosphatase,
DE BF hand calcium-binding domain 1 (PPBF1), transcript variant 1,.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
NCBI_TaxID=9541;
OX NCBI_TaxID=9541;
RN NM_001001.
RP NUCLEOTIDE SEQUENCE.
RA International consortium for macaque cDNA sequencing, analysis;
RT "DNA sequences of macaque genes expressed in brain or testis and its
RT evolutionary implications."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
[2]
NCUBOOTIDE SEQUENCE.
RA Oseada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,
RA Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.,
RA "Substitution rate and structural divergence of 5'UTR evolution:
RT Comparative analysis between human and cynomolgus monkey cDNAs."
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB179074; BAE02125.1; -; mRNA.
SO SEQUENCE 405 AA; 46636 MW; C1B989FP20B2334B CRC64;

```

Query Match 24.4%; Score 839; DB 2; Length 405;
 Best Local Similarity 41.3%; Pred. No. 1e-42;
 Matches 159; Conservative 79; Mismatches 111; Indels 36; Gaps 4;

```

Qy      3 ENATRAAFIOKMYRHRQABRBMORCMQOIFONLEVA5BDOQAEIYKFPNDLIK----- 57
Db      14 ETSLSAALIIQMWYHGYTKLAKQRYALTTFQSLIEYDQSQMLSTFFSFWFNAYAHV 73
Qy      58 HMPQAGRKNOYQGSAAHVLDDKDLVEERGDYVNAKIELPIRNKHIDLDIVERKGRG 117
Db      74 HNGEIELNQSLESQDLRDRHDYVDLDVDPQSYNGRPLQFPLCTCTDIDLLBAKESQO- 132
Qy      118 NRIHKKYALLIIEBAKSKLKQLPNISPVSTAVSQOVTTCGDLHGKLDLLVYTHKGLPS 177
Db      133 -ILAHAYLLEVLFEYKVKYKQMPNPTVQTSKELTICGLDHGLDLDLFLFYNGGLPS 191
Qy      178 SNNPVPFNGDFVDRGKGLVLLLSLYLAFPAVAFNAGNHEDSVNNAARYGFIREVS 237
Db      192 SNNPVPFNGDFVDRGKGLVLLLSLYLAFPAVAFNAGNHEDSVNNAARYGFIREVS 251
Qy      238 KYPRNHRKILAFIDEVRYMLPLGSVTVNSRLVYHGFSDSTSLDLSIDRGKYSILRP 297
Db      252 KYKLGKRIILQLEBELVAMLPITGTVDNELIVHOGISFTTDLNLRHVRNKMGSVLIP 311
Qy      298 PL-----TDGEP--LDKTEWQOIFDIMMSDPQATMGCV 328
Db      312 PLETNRDNDTFSKHNKGVTFNAHRIRKTNQSPSEHLFKHEWQIIDLMSDPKQKQCF 371
Qy      329 PNTLRGAGVWFGPDVTDNFLOHRHL 353
Db      372 PNTGRGGCYGPDVTSKILNKYQL 396

RESULT 17
Q68EP0_XENTR PRELIMINARY; PRT; 493 AA.
ID Q68EP0_XENTR PRELIMINARY; PRT; 493 AA.
AC Q68EP0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Pp5c-prov protein.
GN Name=pp5c-prov;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
NCBI_TaxID=8364;
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RC TISSUS=Embryo;
RX Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RX Strasser R.L., Feingold B.A., Grouse L.H., Derge J.G.,
RX Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RX Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RX Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RX Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RX Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RX Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RX Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RX Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huilyk S.W.,
RX Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RX Fahy J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,
RX Blakesley R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RX Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RX Butlerfield Y.S.N., Krzywinski M.I., Skalska D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
NCUBOOTIDE SEQUENCE.
RA TISSUS=Embryo;
RC TISSUS=Embryo;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +

```

CC phosphate.
 CC -1- SIMILARITY: Belongs to the PPP phosphatase family.
 DR EMBL: BC080162; AA080162.1; -, mRNA.
 DR SNR; Q68EP0; 20-493.
 DR Ensembl; ENSXETG0000018265; Xenopus tropicalis.
 DR GO: GO:0005737; C:cytoplasm; IEA.
 DR GO: GO:0005634; C:nucleus; IEA.
 DR GO: GO:0016787; F:phosphatase activity; IEA.
 DR GO: GO:0004721; F:phosphoprotein phosphatase activity; IEA.
 DR GO: GO:0006470; P:protein amino acid dephosphorylation; IEA.
 DR InterPro: IPR004843; M:pesterase.
 DR InterPro: IPR011236; P:phatase.5.
 DR InterPro: IPR001186; T:phatase_apah.
 DR InterPro: IPR001440; TPR.
 DR InterPro: IPR011990; TPR-like helical.
 DR Pfam: PF00149; Metallophos.1.
 DR Pfam: PF00515; TPR_1/3.
 DR PIRSF: PIRSF03096; PPTase.5/1.
 DR PRINTS: PR00114; STPHPTASE.
 DR Prodom: PD000252; T:phatase_apah; 1.
 DR SMART: SM00156; P2Ac; 1.
 DR SMART: SM00028; TPR; 3.
 DR PROSITE: PS00125; SER_THR_PHOSPHATASE; 1.
 DR PROSITE: PS50005; TPR; 3.
 DR PROSITE: PS50293; TPR_REGION; 1.
 DR Hydrobase: Iron; Repeat; TPR repeat.
 SQ SEQUENCE 493 AA; 56247 MW; DD7597B849DD40EF CRC64;

Query Match 18.2%; Score 627.5; DB 2; Length 493;
 Best Local Similarity 43.8%; Pred. No. 9,4e-30;
 Matches 135; Conservative 48; Mismatches 102; Indels 23; Gaps 8;

QY 105 IDLLIDVFR-KKGRNRLHPYVALILREAAKSLKQLPNISPVSTAVSQVTCGLHGKI 163
 DB 182 IDPMLBIMQFYKQOKLHKRCVQMLVQVKDILSKLPSLVEISLEKSGQVTCGDTNGP 241
 QY 164 DDLIVLHKRGLPSSSNPYFNGDFVDRGKGLGVLLLSLTLAPNNAFLRGNHDS 223
 DB 242 YDLNMTFHLNGLPSENNPFYFNGDFVDRGSPFVEVLTTLFGFLKLVPAQPHLLRGNHED 301
 QY 224 VNAARYGFIRESVEKPRNKRILAFIDEYRWLPGLSVLSRYLVHGGF--SDSTSLD 281
 DB 302 TNNQMGFBEGEVAKY---SAQMFQLESFVQWLPGLAMCNQVLLHNGSLPESDGVTL 358
 QY 282 LKSIDRGKYSILRPLTDGEPDLKTEWQQLFDIMKSDPQATMGCVPTLRGAGVFGF 341
 DB 359 QIRSIENR-----QPF--DSGPM-----CDLLMSDPQPDG--RSSSKRGVSCQFGR 402
 QY 342 DVTDFNPLQRRRLSYVIRSHCKRNGHFMADNKLTITFSASNTYALGSKGAYIRLN-NO 400
 DB 403 DVTRRPLBENGLDYIIRSHVKEGEGVSHNGLCVTVFSAFNPCDQMGNGAYIHLNGSD 462
 QY 401 LMPHFVQY 408
 DB 463 LKPKFHQF 470

RESULT 18
 Q4QGM7 LEIMA PRELIMINARY; PRT; 897 AA.

DT 13-SRP-2005 (TREMBlrel. 31, Created)
 DT 13-SRP-2005 (TREMBlrel. 31, Last sequence update)
 DT 13-SRP-2005 (TREMBlrel. 31, Last annotation update)
 GN Serine/threonine protein phosphatase-like protein (BC 3.1.3.16).
 OS ORFNames=Iamf12.0660;
 OS Leishmania major.
 OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OK NCBI_TaxID=5664;
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Friedlin;
 RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,

RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
 RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
 RA Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL: CT005251; CAJ02710.1; -, Genomic DNA.
 DR Hydrobase.
 SQ SEQUENCE 897 AA; 103525 MW; 16381EB67C70D4F1 CRC64;

Query Match 18.0%; Score 618.5; DB 2; Length 897;
 Best Local Similarity 31.2%; Pred. No. 7,2e-29;
 Matches 190; Conservative 88; Mismatches 215; Indels 115; Gaps 19;

QY 92 VNAKLEPIRKHQIDLVPRKGRNRLHPKVAL--ILREAAKSLKQLPNISPVSTAV 149
 DB 220 VTNLI--PITLHVBOVI---KDIKRVLLPAPMAWFTDAMNHLMTMPNVRVSPAV 273
 QY 150 SQOVT-----VCGDLHGKLDLVLVLRKNGLPSSSNPYFNGDFVDRGKGLGV 198
 DB 274 GARVNGRVIQCKKVVVVDLHGQALDLHLIKBSGMPBSNYYFNGDYVDRGAGVEI 333
 QY 199 LLLLSLYLAFNNAVFLNNGHDSVNAARYGFIRESVEKPRNKRILAFIDEYTRMP 258
 DB 334 ILILFSLMLACPYVTLNNGHCDYNDYGVDFVSTKYDRN--VRLTORCPCALP 390
 QY 259 LGSVLSRVLVHGG--FSDSTSLDLKSIDRGKYSILRPLTDGEPDLKTEWQQLF-D 315
 DB 391 LMTLIGGKIFVHGGPPRRRGVINDISRIQRFQIPI-----PNYSQPEDEIFQD 442
 QY 316 IMMSDPQATMGCVPTLRGAGVFGPDVTDNPLQRRRLSYVIRSHCKRNGHFMADNKI 375
 DB 443 MMSDPVEDIKSMRBSQRGAGVFGSDVMEFLNNKLELIYRSHDCLSGYBEHHNKL 502
 QY 376 ITTFSASNTYALGSKGA--YIRLNQMLMPHFVQY-----ISASQ 414
 DB 503 LTVFSAISNDGSPNSYGAICTEFGNPEPSVHTYQMFEDBYDSQVSLTDSFTLTAANI 562
 QY 415 TKRLSKQRMGYV-----ESSAKKELAVRMDDHRELDEPFXKP 455
 DB 563 ARISSTFATVSGVGGVSKSNFMSRVLLHRRTKDNLRLRRLRIYQRRRLLAYPEKLD 622
 QY 456 KDSGYISISHMKCKMENVTKGLPWRLLRDKLAPGDSQKNV---NRTLDLDVTL 511
 DB 623 TKKGLSMLTEWBSRNVNTLDPYFLRGVLYVRVDHRIYAPFLRPHVLDQ--L 679
 QY 512 EABADGMSYMDALYA---NKAIVAFNIIDANGSEITLDEF--ETAIDLVAHP 563
 DB 680 WGEERROSVCAYVLQOQRNHRSLVSA-----AFNKEVVNVRFCSVRAIDYTT---- 730
 QY 564 GAYSKAEMLKCRMDNDGKVDLNEPLAARLSDLRKEQODENIRRSRGSPVACT 623
 DB 731 ---SDCQLPLESFDEKNGHISGEFL-----KKVOOIAN-----GKDPPLRW 772
 QY 624 ATDPVTL 631
 DB 773 DIDAMBQL 780

RESULT 19
 O42205 XENLA PRELIMINARY; PRT; 492 AA.

DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 GN Protein phosphatase 5 (Fragment).
 OS Name=pps;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OK NCBI_TaxID=8355;
 RP NUCLEOTIDE SEQUENCE.
 RA Ollendorff V., Donoghue D.J.,


```

Db      182 VDMPLMELMGPYKQKGLHRCICLYQMVLQVMDILSKPLSVLVEVLSKSGQVTVCGDTHGCP 241
      164 DDLVLVHLKNGGLPSSSNPVYFNGDFVDRGKGLLEVLILLISLYLAFPNVAVLNRGNHDS 223
      242 YDLMLNIFHLNGLSENNPVIYFNGDFVDRGSPSVSEVITLTFGLYPAHFLILRGNHETD 301
      224 VMNARIGFIREVSKYPRNRKRIILAFIDERYRMPLGSLVNSVTLVHGGF--SDSTSLD 261
      302 TMMQMGFGEVYAKT---SAQMFQLFSVFPQMLPLMVCVNRVLNMGGLFSEIDGVTL 358
      282 LKSIDRGKYSILRPLTDEBPLDKTEWQOIFDIMSDDPQATMGCVPTLFGAGVWFGP 341
      359 QINITERNR-----QRP--DSGPM-----CDLMSDPQPDG-RSTSKRGVSCQFGR 402
      342 DVTDNFLQRRHLSVTRSHCEKPGNGHEFMHNDKLTITFSASNYVALGSKNGAYIRLN-NQ 400
      403 DVTHQFLAEENGLDIYIRSHVEKDEGEVSHNGLCVTVFSAPNYCDQGMNGAYITHSGSD 462
      401 LMPHFVQY 408
      463 LKRFHOF 470

RESULT 21
ID Q7QB71_ANOGA PRELIMINARY; PRT; 474 AA.
AC Q7QB71;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000011234 (Fragment).
GN ORFNames=ENSANG00000008745;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anophelae.
OX NCBI_TaxID=180454;
RN RN
RP NUCLEOTIDE SEQUENCE.
RC The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
RN RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RT Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
RG The Anopheles gambiae Sequence Committee;
RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -1- SIMILARITY: Belongs to the PPP phosphatase family.
DR EMBL; AAA0100880; EAA08659.2; -; Genomic_DNA.
DR GO; GO:0005737; C:Cytoplasm; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004721; F:protein amino acid dephosphorylation; IEA.
DR GO; GO:0006701; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR004843; M-pesterase.
DR InterPro; IPR011236; PPTase_5.
DR InterPro; IPR006186; PPTase_5.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00149; Metallophos; 1.
DR Pfam; PF00515; TPR_1; 2.
DR PIRSF; PIRSF03096; PPTase_5; 1.
DR PRINTS; PR00114; STRPHPTABE.
DR ProDom; PD000252; T_Phtase_ApAn; 1.
DR SMART; SM00156; PP2Ac; 1.
DR SMART; SM00028; TPR; 2.
DR PROSITE; PSS0005; TPR; 2.

```

```

DR PROSITE; PSS0293; TPR_REGION; 1.
KW Hydrolase; Iron; TPR Repeat.
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 474 AA; 54287 MW; A7BDD4FD5BCAC17A CRC64;

Query Match 17.9%; Score 614.5; DB 2; Length 474;
Best Local Similarity 36.4%; Pred. No. 5.5e-29;
Matches 150; Conservative 64; Mismatches 155; Indels 43; Gaps 12;

      11 PLOKWRHRQAREMRQRCNMQIFQNLBYA-----SEDOAELYEKFNLDIKMPQAGR 65
      74 YTKAYRRBAALWALGR--FKALADLEFPAKRCSPAQAQDYTECKKVNKILAFKAI 131
      66 KNOYGSANVSLVD-----KDLVBEFGDIYNAKIELPIRKNHIDLIDVRKRGKR 119
      132 SVQHQKSVKMKCRDMEFATIEDDYGP--KLENGVTTLTFPMK-----LLEWYKKQ--NK 163
      120 LHPKYVALILREAKSLKQLPNTSPVSTAVSQQVTCGDLHGKLDLVLVHLKNGLPSS 179
      184 LHKNFAYRLCDMETLLKTPQSLVEIPVDEQKFTVCGDTHGQFYDLMLFEINGLPSP 243
      180 NPVYFNGDFVDRGKGLLEVLILLISLYLAFPNVAVLNRGNHDSVMNARGYFIREYSKY 239
      244 NPYLFRNGDFVDRGSPVECFITLFGPKLLVPHNFFLAKGNHESFPMNQLYGFGVYVSKY 303
      240 PRNHRRIILAFIDEVYRMLPLGSLVNSRVILVHGGF--SDSTSLDLKSIDRGKYSILRP 297
      304 SQWMAFMFTL---VYNNMLPLCHLIRKKVLYMHGGLFSPKONVSLDDRLSIDRN-----QRP 355
      298 PLTGDGRLDKTEWQOIFDIMSDDPQATMGCVPTLFGAGVWFGPDPVTDNLFQRRHLSYVI 357
      356 P-----EEGLMCELLMSDPHPLPGRVPR--GQVGIQFQPDVTEAFKNNMLDIYII 404
      358 RSHCEKPGNGHEFMHNDKLTITFSASNYVALGSKNGAYIRLN-NQQLMHPHFVQY 408
      405 RSHVEVABGEYVDHNGKCTVFSAPNYCDQGMNGAYITHSGSDLPKFTTY 456

RESULT 22
Q4RPT2_TETNG
ID Q4RPT2_TETNG PRELIMINARY; PRT; 874 AA.
AC Q4RPT2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 16 SCAR15113, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00035185001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN RN
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Micard S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Desliava C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellic V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Catolico L., Poulain J., De Bernardis V.,
RA Cruaud C., Duprat S., Bottelier P., Coutaneau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McMan P., Bosak S.,
RA Kelle M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Lauder V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RN Nature 431:946-957(2004).
[2]

```


Db 653 DKULDGKVSFSPFEQVLTAKOLNINLSABQVRIIM-----RMINSNLSCKKRVQENDKIDVA 708
QY 590 EFLIAR-----LSDLHRK--EQDENTIRRS 614
Db 709 EFIGKRVCRALINKYINNEKVOKLITIGKHLADSEIANFHYRFEYEQDSENAUR 768
QY 615 TGRPSVAKT 623
Db 769 RKRSVYKS 777

RESULT 24
081728 TRYCR
ID 081728 TRYCR PRELIMINARY; PRT; 923 AA.
AC 081728;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE TCC31.31.
GN Name-TCC31.31;
OS Trypanosoma cruzi.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma;
OC Schizotrypanum.
OX NCBI_TaxId=5693;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Brener;
RX MEDLINE=98391765; PubMed=9724326;
RA Anderson B., Aelund L., Tammi M., Tran A.N., Hohelsel J.D.,
RA Pettersson U.;
RT "Complete sequence of a 93.4-kb contig from chromosome 3 of
RT Trypanosoma cruzi containing a strand-switch region.";
RL Genome Res. 8:809-816(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Brener;
RA Anderson B., Bontempi E.J.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL, AC137988; AAN78342.1; -, Genomic_DNA.
DR HSSP, P36873; 1176.
DR GO, GO:0016787; P:hydrolase activity; IEA.
DR InterPro, IPR011992; BF-Hand type.
DR InterPro, IPR002048; EF_hand_Ca_bd.
DR InterPro, IPR004843; W-pesterase.
DR InterPro, IPR006186; T-phlase_apaH.
DR Pfam, PF00149; Metallophos; 1.
DR PRINTS, PR00114; STPHPTAS.
DR ProDom, PD000252; T.phlase_apaH; 1.
DR SMART, SM00156; P2AC; 1.
DR PROSITE, PS00125; SER_THR_PHOSPHATASE; UNKNOWN_1.
SQ SEQUENCE 923 AA; 104798 MW; 95B6646B32B2143 CRC64;

Query Match 17.3%; Score 594.5; DB 2; Length 923;
Best Local Similarity 33.0%; Pred. No.2.1e-27;
Matches 174; Conservative 73; Mismatches 206; Indels 75; Gaps 15;

QY 128 IIREAKSLKQLEPNI-----SPVSTAVS-----QQVTCGDHGLKLDLLVLAHKNGLP 176
Db 299 VFIDASHLNTMNVVRLSPVGVARVNGRINSKVVVGDHGLADLLHLTKCGMP 358
QY 177 SSSNPVFNQDFVDKRGGLLEVLILLISLYLAPNAVFLNRGNHEDSVNARVGFIREY 236
Db 359 NESTYVIFNGDFVDKRGAGVEVILLIFSLMLACPKYTLNRGNHEDCVNMDDYGPVENS 418
QY 237 SKYPRNHRKLIATIDEVYRWLPGLSVLNSRVLLVHGPF--SDSTSLDLTKSIDRGKVS 294
Db 419 TKIDRN---VRLVQRCFLPLATITIGKVFVHGLPRKGVNIEDISRIQRFROI-- 473
QY 295 LRPEPLDGEPLDTEWQOIF--DIMSDPQATMGCVPTLNRGAGVWFGPDVDTNPLQRHRL 353
Db 474 ---PMPD---YSQPEEDIEIFODLLMSDPVEDLQGMBSRPGAGVVGADVTOGFLQNGL 527
QY 354 SYVIRSHCEKPNHGFPHNDKITTIPSASNYVAIGSKAV--IRLNQMLPHFVQYISAA 412

Db 528 ELVIRSHCEACGYEBEHHGKLLTVFSASNYOGPEINPFSFPAVVDNNEPSHTYQVAE 587
QY 413 SOTK-----RLSPKRMGIVESA-----LKLAVMRDRHDELD 448
Db 588 DVEVQTLVDLGETFPTLGRVSPFATLSQSKLLRRARADVLRVLRRIYGRRLRLA 647
QY 449 EPRKIDPKDSGYTISISHCKVWENVTYKGLPRLRLDKLAPGDSQKATYNNRL----- 502
Db 648 YPAKLDRTKSGSVWKLEWYEAARMVLTLDLPWFPLRGYLVADEENTRWVSHFLVFNH 707
QY 503 --DLDTVILREADGMSVMDALYANKASLVAIFNIIDANSGBITLLEPFAIDLLVA 560
Db 708 FOPLMLNDWIOSH---CHHLTOOQRANHRQY---VAQAFNBOVSINPESVIRAIIDY 760
QY 561 HNPAGYAKMLEKRCMDLNGDKYDLNEF-----LEAFRLSLDHR 602
Db 761 TW-----SEAQLQLFVYLDGGTGHDGKPFVNMLEMAAYPLSDTLR 804

RESULT 25
0512P5 ENTHI
ID 0512P5 ENTHI PRELIMINARY; PRT; 473 AA.
AC 0512P5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Protein phosphatase, putative.
GN ORFName=90.t00028;
OS Entamoeba histolytica HM-1:IMSS.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxId=294381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HM-1:IMSS;
RX PubMed=15729342; DOI=10.1038/nature03291;
RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
RA Amedeo P., Roncaglia P., Bettman M., Hitt R.P., Mann B.J., Nozaki T.,
RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leippe M.,
RA Hofer M., Bruchhaus I., Willhoeft U., Bhattacharya A.,
RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
RA Quail M.A., Rabinowitsch E., Norbertczak H., Price C., Wang Z.,
RA Guillen N., Gluchist C., Strop S.E., Bhattacharya S., Iohia A.,
RA Foster P.G., Sichteritz-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrett B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite Entamoeba histolytica.";
RL Nature 433:865-868(2005).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL, AAPB01000329; BALT4840.1; -, Genomic DNA.
SQ SEQUENCE 473 AA; 53606 MW; 665B460BC0ED93A8 CRC64;

Query Match 16.9%; Score 582.5; DB 2; Length 473;
Best Local Similarity 35.4%; Pred. No.4.8e-27;
Matches 143; Conservative 74; Mismatches 134; Indels 53; Gaps 14;

QY 11 FIOKTYRRHQ-----REMRRCWQIFQNLIVSBDQAEYLFYFDLIGHMQAAG 64
Db 68 FVGYGRQASAPAAQLQLEAISAAC-----EAKKSLSPD-----GMINSNGLKREK-- 115
QY 65 RKNQYOGSAHVSGLDDKDL-----VEFGDIYNAKIELPRKNHIDLLIDVPRKGN-- 118
Db 116 RREQLPFEA-ISEVBEKQITISWRDIDASTSTIKIEBQKITENYQ--EIVAMSNQ 171
QY 119 ---RLHPTVALLIREAKSLKQLEPNI-----SPVSTAVSQQVTCGDHGLKLDLLVLAHKN 174
Db 172 YGKCAHLTKYCRITIEBAGEIILKRVNLTLEINTQ--GKRWTVGDHIGCFPDILHLEING 230
QY 175 LPSSNPVFNQDFVDKRGGLLEVLILLISLYLAPNAVFLNRGNHEDSVNARVGFIRE 234

| | | | |
|--|--|--|-----|
| Dd | | 231 LPSEDNVYLTFNQDPVDKRGFGVCULTLFEPMIIPYNSFLFARGNHETPANAAMATGFSGE | 290 |
| Oy | | 235 VSKRPPRNKRRIIAAIFDEYYRWLPGLSVINRYLIWGGFPSDS-TSLDLIKSIDRKGYVS | 293 |
| Dd | | 291 VKTKY--NENNYALAFSD-VFMOLPYCHVIDSKIYVVHGIIPPMTYSILDIDKKIKKG---- | 343 |
| Oy | | 294 ILRPLTQSEPDKNKEMOQIIPIIMSGDPAQTGCVPNTIRAGGWGFPGPVTDNFLQRHL | 353 |
| Dd | | 344 -----CDPOGSIALSALLMWADPTSNSSPS-STRGCKSGSPDTITNFILGNWL | 391 |
| Oy | | 354 SVYIRSHCEKPNGHEHFMDKNKIITTFISAGNYVALGSNKAYTRL | 397 |
| Dd | | 392 QIVRSHEMKRMNGYEWGAEGRRLITYSAANYCDDQMNNKAAYIHV | 435 |
| RESULT 26 QNES8 CAEEL | | | |
| ID | QNES8 CAEEL PRELIMINARY; | PRT; 496 AA. | |
| AC | QNES8: | | |
| Dt | 01-OCT-2000 (TRYMBLrel. 15, | Created) | |
| Dt | 01-OCT-2000 (TRYMBLrel. 15, | Last sequence update) | |
| Dt | 01-MAR-2004 (TRYMBLrel. 26, | Last annotation update) | |
| Dd | Hypothetical protein yj39B6A.2. | | |
| GN | ORFNAMES=yj39B6A.2; | | |
| OS | Caeornhabditis elegans. | | |
| OC | Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditoidea; | | |
| CC | Rhabditidae; Felodetrinae; Caenorhabditis. | | |
| OK | NCHI_TaxId=6239; | | |
| RM | [1] | | |
| RP | NUCLEOTIDE SEQUENCE [LONG SCALE GENOMIC DNA]. | | |
| RC | STRAIN-Bristol NZ: | | |
| RX | MEDLINE=99069613; PubMed=9851916; | | |
| RG | The C. elegans sequencing consortium; | | |
| RT | "Genome sequence of the nematode C. elegans: a platform for | | |
| RT | investigating biology."; | | |
| RL | Science 282:2012-2018(1998). | | |
| DR | EMBL; AL132948; CACS1076.2; -; Genomic_DNA. | | |
| DR | PIR; T45058; T45058. | | |
| DR | HSP; P53041; LA17. | | |
| DR | Infect; QNES8; - | | |
| DR | Ensembl; YJ39B6A.2; Caenorhabditis elegans. | | |
| DR | Wormbase; WBGene0012665; YJ39B6A.2. | | |
| DR | WormPept; YJ39B6A.2; CR36619. | | |
| DR | GO; GO:0005737; Cytoplasm; IEA. | | |
| DR | GO; GO:0005634; Cilium; IEA. | | |
| DR | GO; GO:0005488; F-binding; IEA. | | |
| DR | GO; GO:0016787; Phosphodiesterase activity; IEA. | | |
| DR | GO; GO:0004721; Phosphoprotein phosphatase activity; IEA. | | |
| DR | GO; GO:000470; Protein amino acid dephosphorylation; IEA. | | |
| DR | InterPro; IPRO014643; M-pesterase. | | |
| DR | InterPro; IPRO11236; PPase_5. | | |
| DR | InterPro; IPR006186; T_phctase_apah. | | |
| DR | InterPro; IPR001440; TPR. | | |
| DR | InterPro; IPRO11990; TPR-like helical. | | |
| DR | pfam; PF00149; Metallophos; 1. | | |
| DR | pfam; PR00515; TPR 1; 2. | | |
| DR | PIRSF; PIRSFP03096; PPcase_5; 1. | | |
| DR | PRINTS; PRO0114; STPHPTSASE. | | |
| DR | Prodrom; PD000252; T_phctase_apah; 1. | | |
| DR | SMART; SMO0028; TPR; 2. | | |
| DR | PROSITE; PS50005; TPR; 2. | | |
| DR | PROSITE; PS50293; TPR_REGION; 1. | | |
| KM | Complete proteome; Hypothetical protein. | | |
| SQ | SEQUENCE 496 AA; 56462 MW; 2CB872FF1700384EE CRC64; | | |
| .Query Match 16.9%; Score 581; DB 2; Length 496; Best Local Similarity 33.0%; Pred. No. 6,3e-27; Matches 137; Conservative 77; Mismatches 167; Indels 34; Gaps 12 | | | |
| Oy | 3 ENAIRAI---FIQMYRHQAAREMR-RCMNIQFONTEVASBODAEIYVF--ENDII 56 | | |
| Dd | 86 DNAI--AIPSYVKYGRRATANMALGRFKALTDOAVLVKCPCPNKDARAADFECskiv 143 | | |

[illegible]

RESULT 27
Q42553 PLABE PRELIMINARY; PRT; 829 AA.
AC Q42553;
DT 13-SEP-2005 (TRMBLrel. 31, Created)
DT 13-SEP-2005 (TRMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRMBLrel. 31, Last annotation update)
DE pp1-like protein serine/threonine phosphatase, putative
DS (fragment).
GN ORFNames=PB000626.00.0/
OS Plasmodium berghei.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxId=5621,
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karras M., Raine J.D., Carlton J.M., Kool J.T.W.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quill M.A., Ormond D., Doggett J., Trueman H.B., Mendoza J.,
RA Bitwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos
RA Jansse C.J., Barrett B., Turner C.M.R., Waters A.P., Stiden R.S.
RT "A comprehensive survey of the Plasmodium life cycle by genomic
RT transcriptomic, and proteomic analyses.";
RL Science 307:82-86(2005).
-1 CAUTION: The sequence shown here is derived from an
EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -1 CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1 SIMILARITY: Belongs to the PPP phosphatase family.
CC EMBL: CAAI01000476; CH9448.1; -; Genomic_DNA.
DR InterPro: IPR011993; EF_Hand_type.
DR InterPro: IPR002046; EF_Hand_Ca_Bd.
DR InterPro: IPR004843; M-pesterase.
DR InterPro: IPR006186; T_phase_apah.
DR Pfam: PF00036; ehand; 2.
DR Pfam: PF00149; Metallophos; 1.
DR PRINTS: PRC0114; SRPHPHAS.
DR PRODOM: PD000252; T_phase_apah; 1.
DR SMART: SM00054; EFn; 3.
DR SMART: SM00156; PPAC; 1.
DR PROSITE: PS00125; SER_THR_PHOSPHATASE; UNKNOWN_1.
KW Hydroxylase; Iron.
FT NON TER 1
SQ SEQUENCE 829 AA; 96742 MW; 87CD24D615A8FC7 CRC64;
Query Match 16.84; Score 578; DB 2; Length 829;
Best Local Similarity 27.44; Pred. No. 1.9e-26;

| Query Match | 16.8% | Score 577.5 | DB 2 | Length 594 |
|---|-----------------|-------------------|-----------|------------|
| Best Local Similarity | 35.9% | Pred. No. 1.3e-26 | | |
| Matches 151 | Conservative 77 | Mismatches 144 | Indels 49 | Gaps 17 |
| DR InterPro: IPR004843; M-pesterase. | | | | |
| DR InterPro: IPR011236; ppIase_5. | | | | |
| DR InterPro: IPR001440; TPR. | | | | |
| DR InterPro: IPR006186; T_phtase_apah. | | | | |
| DR Pfam: PF00149; Metallophos_1. | | | | |
| DR Pfam: PF00515; TPR_1, 3. | | | | |
| DR PIRSF: PIRSF03096; pPtease_5, 1. | | | | |
| DR PRINTS: PR00114; STPHPTASE. | | | | |
| DR SMART: SM00156; PP2AC; 1. | | | | |
| DR SMART: SM00028; TPR; 3. | | | | |
| DR PROSITE: PS50005; TPR; 3. | | | | |
| DR PROSITE: PS50293; TPR REGION; 2. | | | | |
| SQ SEQUENCE 594 AA; 69196 MW; 63E3E3750664181C CRC64; | | | | |
| QY 6 IRAAIFIOKWYRRHQAREMQRRCNMQI PONTLEVAS--ODQAELYKFPNDLIKMPQAA 63 | | | | |
| DB 184 IDEAIKINPY-----AKAYRKGSYLLSDLRASBCFOGVLTSLTKDKNSELK-LKQCK 238 | | | | |
| QY 64 GR--KNQYQSAHYSLVDDK---DDLVERGDIIVAKIELPT-RKQHIDLLDVERKKR 116 | | | | |
| DB 239 KLIFEOGFQKAI--LEQKKPYETLVLSLKEN--MEAPYIDRNNTLV--DFLKVA 291 | | | | |
| QY 117 -----GNRLAPKYVALILREAAKSLKQLPISPSTAVASQOVTGCDLHGKLDLLVVL 170 | | | | |
| DB 292 DYISIPNNKMLKKCVCAIVLDVITKLKSLPLVYLINLEBEDTTLTGCVHQYYDLNIM 351 | | | | |
| QY 171 HKNGLPSSSNPYVNGDFVDRGKGLSVLLLSLYLAFPNAVFLNGNHDSVMNARYG 230 | | | | |
| DB 352 KINGPSPKNSYLPFGNDPVDGRSPSEVITFLYLAKLTFPNNVYLTGNHETDNMTYIG 411 | | | | |
| QY 231 FIREVSEKTPYNNKRIILAIFDEVYRMTPLGSLVNSRVLYHGEF--SDSTSLDLIKSIDR 288 | | | | |
| DB 412 FLGLGLOEKY--DEKMHVLFSDS-FKPLPLAYVLNKNIFICHGGIPSKTDTTLBDEIKRIDR 468 | | | | |
| QY 289 GKVYSILAPRLTDLDEPLDPTKEWQOIPIDMMSDPAITMGCVPNLTRGAGWPGPDVYDNL 348 | | | | |
| DB 469 NK-----EPDSE--GVMTLMSDPEERKGFPSK-RGIGSPGDTITENFL 512 | | | | |
| QY 349 QRHRLSYVIRSHCEKPGNGHERMDNKKITITIFSASNYAIGSNKAYIRL-NNQMLPFPVQ 407 | | | | |
| DB 513 KINMLSLIRSHYVRDDEGSLGQNGQLYTFPSAPYDCDKNNKGAFLFKGNSIKPECVT 572 | | | | |
| QY 408 Y 408 | | | | |
| DB 573 F 573 | | | | |


```

Oy 50 ---KFFNDLLKHPPOAAGRNQYQGAHVSVLDDODLVEERGDVNAKI-----ELP 99
Db 149 DRSLKLTHTETDPNAPGFAVSNDSYG--PHKXPLVDHSKTAQGS-GNETNKSSEFVTPVPD 206
Oy 100 ---IRKNHIDLIDVERKKRGNRLHPKRYALLIRBAASKLQOLPIBSFVPAVSQO--VT 154
Db 207 ASLIRBSFVTEILD-FLKNPBNRLHKKVAYMYLVODLQVLKEVASKPLVRINIOKSHIT 255
Oy 155 VCGDLDJGKLDLIDLVTLHKGLPSSSNPYYFNQDFVDGRGKLEVLILLISLYLAPNAVP 214
Db 266 VCGDHQGFDDLNTINIFINGLPSVNNNGYLFNQDFVDRGSFSEVYLIVFTLKIWPYHVH 325
Oy 215 LNRGNHDSVMNARYGPIREBSKPYRNRKRLIARLIDVYKRLPGASTLNRVLVYNGCF 274
Db 326 LARGNHETKQNLKLYGBGSEVLAKDSG--LYDLISBAFCFLPLAHVINDKVVVYHGL 382
Oy 275 SDSTSLDLIKSIDRGKVSILRPPLTLDGSEPLDKTEMOQIFDIIMSDPOATWGCVPNLRG 334
Db 383 CSEBNVYL-----SDIBLYSRCEPADS---GMSSTIIMSDPOQKGRSPSP-RG 428
Oy 335 AGVWFGEVDVTNFIQHRRLSYVIRSHCKPNQGEHMAHDKNIITIFSASNYVALGSNGAY 394
Db 429 VGCNFGPDVTLNFIKTNINLDYIIRSHLEVYQBGYSVDHDKCIITYBSAPVCDSGMKNGAF 488
Oy 395 IRLAN-QLMHPFOY 408
Db 489 IKIHEYDLKENFVOF 503

```

| RESULT 32 | Q4UPF6_THEAN | PRT: | 516 AA. |
|-----------------------|---|---|-------------------|
| AC | Q4UPF6_ | THEAN PRELIMINARY; | |
| DT | 13-SEP-2005 | (TREMBLrel. 31, Created) | |
| DT | 13-SEP-2005 | (TREMBLrel. 31, Last sequence update) | |
| DT | 13-SEP-2005 | (TREMBLrel. 31, Last annotation update) | |
| DE | Serine/threonine protein phosphatase, putative (EC 3.1.3.16). | | |
| GN | ORFNames=TH1575; | | |
| OS | Thellieria annulata. | | |
| OC | Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Thellieridae; | | |
| OC | Thelliera. | | |
| OX | NCBI_TaxId=5874; | | |
| RN | [1] | | |
| RP | NUCLEOTIDE SEQUENCE. | | |
| RC | STRIN=Amkra isolate clone C9; | | |
| RA | Pain A., Renauld H., Murphy L., Harris D.A., Quail M.A., Berriman M., | | |
| RA | Hall N., Barreil B.G.; | | |
| RT | "The chromosome 2 sequence of Thellieria annulata." | | |
| RL | Submitted (Apr-2005) to the EMBL/Genbank/DBJ databases. | | |
| DR | EMBL; CR940348; CA174040.1; -; Genomic_DNA. | | |
| KM | Hydrolase. | | |
| SO | SEQUENCE 516 AA; 59193 MW; C670662EBC4FA55 CRC64; | | |
| Query Match | 16.7%; | Score 574.5; | DB 2: Length 516; |
| Best Local Similarity | 38.2%; | Pred. No. 1.6e-26; | |
| Matches 120; | Conservative 65; | Mismatches 102; | Indels 27; Gaps 8 |
| QY | 103 NHID--LLIDV--FRKKRGNRLHPRTVALILREAKSLKQDNPISPVSTAVSQOVTGCG 157 | | |
| DB | 200 NHIDSYVFNHNTLEVIKVPNGNTIHKKVCMIILEVIKIRREVSVDIMNIPDELTGCG 259 | | |
| QY | 158 DLHGGLDLVLVHLNKGLPSSNPVFNDDPYDRGRGLEGVLLLLSLYLAIPNVAFLNR 217 | | |
| DB | 260 DIHGQPYDLMLNPSINGEPTDINSYFNDDPVDKRSFSFECVTLFLAVLVLPSSPHIVR 319 | | |
| QY | 218 GNHSDVNNARQYFTIREVSKKPYRHKRIALFIDEVYRMLPGVSYNSHVLIVHGPF--S 275 | | |
| DB | 320 GNHFEALANKKCGFGEILNKKY--DEKYNVNFCSBFRLPLGILNKKVVLVHGLFGT 376 | | |
| QY | 276 DSTSLDLIKSIDRGKYSILRPLLDGEPDLDTHEMOQIFDIMSDFQATMGCVPTNLRGA 335 | | |
| DB | 377 ENVTIEDLKKIDRFK-----EPSSD--GLMTDMIMSDPKPENGIGLSPSK-RCV 420 | | |

```

QY      336  GWA6RGD1VD1ND1FLQ1RR1LS1VR1SH1CK1DP1NG1HE1FM1DN1K1IT1TS1AS1NY1VA1IS1GN1KAY1 396
Db      421  GFERGPD1IS1FS1KN1LS1Y1TR1SH1EV1KQ1B1GE1VE1HD1K1Y1IT1SA1PN1Y1CD1GM1KA1FI1 480
QY      396  RL1INQLMP1HE1VO1Y 408
Db      481  RIKGDTLE1PR1FT1OP 494

```

| RESULT | 33 | |
|--------|---|---|
| ID | Q962N7_PLAFA | PRELIMINARY |
| AC | Q962N7 | PR: 594 AA. |
| DT | 01-DEC-2001 | (TREMBLrel. 19, Created) |
| DT | 01-DEC-2001 | (TREMBLrel. 19, Last sequence update) |
| DT | 01-MAR-2004 | (TREMBLrel. 26, Last annotation update) |
| DE | Serine/threonine protein phosphatase PfPP5. | |
| GN | Name=pp5; | |
| OS | Plasmodium falciparum. | |
| OC | Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium. | |
| NCBI | NCBI_TaxID=5833; | |
| RM | [1] | |
| RP | NUCLEOTIDE SEQUENCE. | |
| RX | MEDLINE=21896223; PubMed=11897131; DOI=10.1016/S0166-6851(02)00007-5; | |
| RA | Lindenhal C., Klinkert M.Q.; | |
| RT | "Identification and biochemical characterization of a protein | |
| RT | phosphatase 5 homologue from Plasmodium falciparum."; | |
| RL | Mol. Biochem. Parasitol. 120:257-268(2002). | |
| DR | EMBL, AF0404815, AAK95648.1; -; Genomic_DNA. | |
| DR | HSPF, P36873; 1176. | |
| DR | GO; GO:0005737; C:cytoplasm; IEA. | |
| DR | GO; GO:0005634; C:nucleus; IEA. | |
| DR | GO; GO:0016787; F:hydrolase activity; IEA. | |
| DR | GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA. | |
| DR | GO; GO:0006470; P:protein amino acid dephosphorylation; IEA. | |
| DR | InterPro; IPR004843; M-peptidase. | |
| DR | InterPro; IPR011236; pp2ase_5. | |
| DR | InterPro; IPR001440; TPR. | |
| DR | InterPro; IPR006186; T.phase_apha. | |
| DR | Pfam; PF00149; Metallophos; 1. | |
| DR | Pfam; PF00515; TPR_1; 3. | |
| DR | PIRSP; PIRSF033096; pp2ase_5; 1. | |
| DR | PRINTS; PR00114; STPHPTASE. | |
| DR | SMART; SMO0156; PP2Ac; 1. | |
| DR | SMART; SMO0028; TPR; 3. | |
| DR | PROSITE; PSS0005; TPR; 3. | |
| DR | PROSITE; PSS0293; TPR_REGION; 2. | |
| SO | SEQUENCE 594 AA, 69248 MW, 64CA9D0004DEEEF CRC64; | |

| | | | | |
|-----------------------|-----------------|--|-----------|------------|
| Query Match | 16.7% | Score 573.5 | DB 2 | Length 594 |
| Best Local Similarity | 35.9% | Pred. No. 2.3e-261 | | |
| Matches 151 | Conservative 77 | Mismatches 144 | Indels 49 | Gaps 17 |
| QY | 6 | IRAAIFIDKWRRIHQARREMRQRCMIOIFOMLEVASE--ODQABLYKFENDILKIMPOAA | 63 | |
| DB | 184 | IDBAIKIMIPY----AKAYRRKGSGYLLSLDKRASBCFOYKLTQKDNSELK-LKQCK | 238 | |
| QY | 64 | GR--KNQYQSAHVSVDK----DLYBSPDIDYNAKIELPI-RKNHIDLIDVPRKR | 116 | |
| DB | 239 | KLIFQOQFOKALB--LEQKMPYETTLVDSLKIKN--MEATYRRNNTL--DLKCKAA | 291 | |
| QY | 117 | -----GNRLHPKYVALILIRBAKSLKOLPNISPYSTAVSQQVTCGDLHGKLDLIVL | 170 | |
| DB | 292 | DYISFNNKLNKKCYCALIVLVITKLKELPTLVYINTLEDEFTLTCGVHQGYUDDLNTM | 351 | |
| QY | 171 | HNQGLPSSNRYVFNFGDVPDRGKGLKELVLLLSLYLAFPNAVFLNRRGNHDSVNNARYG | 230 | |
| DB | 352 | KINGIPSSKNSYLLTFNGDFVDRGSEVEVYIIFFLYLAKLFPNNVYLIRGHRJEDNNMKIYG | 411 | |
| QY | 231 | FIREVASKYPPNNHKRIILAFIDEVYKMDPLGSLVNSRVLIVHGGF--SDSTSLDLIKSIDR | 288 | |
| DB | 412 | FLGELQEKCY--DEKHNHVFSDS-FEFLPLAYLVNNTNIFLCHGIGISKDTTTLIEDIEKIR | 468 | |

```

OY 289 GKYVSLNPLPLDGEPLKTEWQOQ.FD1MMSPDQATMGVUPPTLRAGAGWREGPDYDNEFL 348
DB 469 IK-----EPLDB--GVMTDLTMSDNEPKPEK-8GIGSFSTDTITENFL 512
OY 349 OPHRLSYVIRSRCEKPNCHGEFHMDDKIIITFSPASNYVAIGSKGAYIRL-NNQLMPHFVQ 407
DB 513 KINNLSLIRSHVADDEGVSLQNGQITVTFSPAIPYCDIMNKGAFLFKGNSITDECVT 572
OY 408 Y 408
DB 573 F 573

RESULT 34
OSW6J3_ORYSA
ID OSW6J3_ORYSA PRELIMINARY; PRT; 483 AA.
AC OSW6J3;
DT 01-FEB-2005 (TREMBLrel. 29, Created)
DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE Putative serine/threonine phosphatase.
GN Name=OSUNB0115P21.1;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Brhatrioidae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL "Oryza sativa BAC OSUNB0115P21 genomic sequence."
RT Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR BMBJ; AC135918; AAV4139.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR004843; M-pesterase.
DR InterPro; IPR011236; Peptase_5.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR006186; T_phase_apah.
DR Pfam; PF00149; Metallopinos; I.
DR Pfam; PF00515; TPR_1; 3.
DR PIRSF; PIRSF033096; Peptase_5; 1.
DR PRINTS; PR00114; STEPHPTASE.
DR Prodom; PD000252; T_phase_apah; 1.
DR SMART; SM00156; PP2Ac; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PSS0005; TPR; 3.
DR PROSITE; PSS0293; TPR_REGION; 1.
SQ SEQUENCE 483 AA; 54448 MW; C652EDB87A1A44DA2 CRC64;

Query Match 16.6%; Score 570.5; DB 2; Length 483;
Best Local Similarity 34.8%; Pred. No. 2.6e-26;
Matches 145; Conservative 69; Mismatches 154; Indels 49; Gaps 122

OY 9 AIFDKWRBRHQARERMQR-RCNMQIPONLVEYASQDDQAELYKFNDDLIKHMPOAAGRN 67
DB 77 ARYSKGYRRGAAYLAWGKFKKALXDFQVQKRIISPNDP-----DATTRKUKCEKAVQ 128
OY 68 OYQGSANHSVLDDKDDLVEEFGDINAKIETL-----PIRKNHIDLIDVPRKRG 117
DB 129 KIRFSEBALSVGBEKGKSVADSIDYRIIIEVEPQYTPRVDGDTITIDPYKAMLDEFEKK- 187
OY 118 NLLAPRYVALIIRBAKSLKQLPNTISPVSTAVSQOVTYCGDILHGKLDLLVLRHNGPS 177
DB 188 -CINRKYAVQVTLQTVQLRSVPSLVVDVWPPDGSHFTVCGVDHGQYFDLLNFKLNGPS 246

```

| | | | |
|---------------|---|---|--------------|
| QY | 178 | SSNPVYVGGDPRDREKRGLEVLILLLSLYLARPNVAFLNRGHEDSVNNARGFIEVES | 237 |
| DB | 247 | ENNPFLFNFGDPRDGRGSFVSVELTLTPAFKCLVPTGVIYLARGNHESKNNKLTGFGEVRS | 306 |
| QY | 238 | KYPRNRHKILAFID---EYWRVLPGLSVLNSHYLVTHGQ-FS-DSTSLDLKSIDRGKYV | 2922 |
| DB | 307 | KLGE-----AFIELFAERFCCLPFLAHYINNVFVHGGLFSDVGKLDIAIDR---- | 356 |
| QY | 293 | SILRPPLTGDREPLDXTENMOQIFDLWMSDPQATMGCVPTLTGAGWGFGDYDNLQORR | 3522 |
| DB | 357 | --FREPPREG-----LMCEVLMSDPQPOLQGRGSPK-RGVGLSFGADVTKCFIQENN | 4040 |
| QY | 353 | LSYVIRSRCKCNHGFPHNDKLTITFSASNYALGNNGAYIRLN-NOLMPFVQY | 408 |
| DB | 405 | LDLVVRSHVDEGGEIIEHDGKLTIVFSAPNTCDQNGNKAYIRFTAPELKPISSF | 461 |
| RESULT 35 | | | |
| QBHIH4_LYCSES | | | |
| ID | QBHIH4_LYCSES | PRELIMINARY; | PRT; 465 AA. |
| AC | QBHIH4; | | |
| DT | 01-MAR-2003 | (TREMBLrel. 23, Created) | |
| DT | 01-MAR-2003 | (TREMBLrel. 23, Last sequence update) | |
| DT | 01-FEB-2005 | (TREMBLrel. 29, Last annotation update) | |
| DE | Type 5 serine/threonine phosphatase 55 kDa isoform (Type 5 protein | | |
| DE | serine/threonine phosphatase 55 kDa isoform). | | |
| GN | Name=PP5; | | |
| OS | Lycopersicon esculentum (Tomato). | | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | |
| OC | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asteridae; | | |
| OC | Lamiid; Solanales; Solanaceae; Solanum; Lycopersicon. | | |
| OX | NCBI_taxonomy:4081; | | |
| RN | (1) | | |
| RP | NUCLEOTIDE SEQUENCE. | | |
| RX | PubMed:12972652; DOI=10.1104/PP.103.02617; | | |
| RA | De la Fuente Van Bentem S., Voosen J.H., Voosen J.B., | | |
| RA | de Vroemen M.J., Gadella T.W.J. Jr., Haring M.A., Cornelissen B.J.C.; | | |
| RL | Submitted (NOV-2002) to the EMBL/Genbank/DBS databases. | | |
| DR | EMBL; AY150041; AAC64317.1; -; mRNA. | | |
| DR | EMBL; AY182778; AAC62214.1; -; Genomic_DNA. | | |
| DR | HSSP; P53041; IAI7. | | |
| DR | GO; GO:0005737; C:cytoplasm; IEA. | | |
| DR | GO; GO:0005634; C:nucleus; IEA. | | |
| DR | GO; GO:0005486; F:binding; IEA. | | |
| DR | GO; GO:0016787; F:hydrolase activity; IEA. | | |
| DR | GO; GO:0004629; F:phospholipase C activity; IEA. | | |
| DR | GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA. | | |
| DR | GO; GO:0004724; P:intracellular signaling cascade; IEA. | | |
| DR | GO; GO:0006470; P:protein amino acid dephosphorylation; IEA. | | |
| DR | InterPro; IPR004843; M-pesterase. | | |
| DR | InterPro; IPR000909; PI_PLC_X. | | |
| DR | InterPro; IPR011236; pp2ase_5. | | |
| DR | InterPro; IPR011990; TPR-like_helical. | | |
| DR | InterPro; IPR006186; T_phase_apah. | | |
| DR | Pfam; PF00149; Metallophos; 1. | | |
| DR | Pfam; PF00515; TPR_1; 3. | | |
| DR | PIRSF; PIRSF033096; pp2ase_5; 1. | | |
| DR | PRINTS; PR00114; STPHPRASZ. | | |
| DR | ProDom; PD000252; T_phase_apah; 1. | | |
| DR | SMART; SM00156; PP2Ac; 1. | | |
| DR | SMART; SM00028; TPR; 3. | | |
| DR | PROSITE; PS50007; PIPLC_X_DOMAIN; 1. | | |
| DR | PROSITE; PS50005; TPR; 3. | | |
| DR | PROSITE; PS50293; TPR_REGION; 1. | | |

SEQUENCE 485 AA; 54720 MW; FCF68708D5520B68 CRC64;
Query Match 16.6%; Score 570.5; DB 2; Length 485;
Best Local Similarity 36.1%; Pred. No. 2.6e-26;
Matches 140; Conservative 55; Mismatches 124; Indels 69; Gaps 12;
QY 24 EMQRRCNMQIFQNLLEYASEDDQALYKFENDLIKHPQAGRNQVQGSANHSVLDKDD 83
DB 142 EGRRR---SVADSIDRS-----VVEPQIAGARIE----- 169
QY 84 LVEEFQDIYNAKLELPFRKNHIDLIDVPRKRGNRLHPFYVALILREAKSLKQLPNIS 143
DB 170 -----GDVVTLL-----DFVKMLDDFRNQK--NLHKRYAVQYILQTRHMLRALPSLV 214
QY 144 PUSTANVQGVTVCGDLHGKLDLLVYLHKNGLPSSNPVYFENDFDPRGKGLFVLLLL 203
DB 215 DIVPBGKFTYVGGDVHGFYDILNLFELNGLPSEBNPVLFNDFDPRGSPSLFVLTITF 274
QY 204 SYLIAPNNAVFLNRGNHDSVNNARVGFIREVSKYPRNHRKILAFIDEVYRMLPLGSYL 263
DB 275 AFRCMCPSAHLARGNHESKMKITGFEGEVASKLSIFVELFA---EVCCLPLAHYI 331
QY 264 NSRVLIYVHG--FS--DSTSLDLKSIDRGKYSILRPPLTDEBPLDKTEWQQLFDIWMSP 321
DB 332 NEKVFYVHGGLFSEVDGKLSIDIRAIRD-----FCBPP-----EKGIMCEILWSDP 376
QY 322 QATMGCVPTLRAGVFGPDVTDNPLQHRRLSYVLRSHCKRNGHEFMHDKITITTEA 381
DB 377 QPQPGGPFK--KGVGLSPGSDVYKRLQERNLIDLVRSHSVDEGVIEIRHDKLITVFA 435
QY 382 SNYYAIGSNKAYIRLN--NQLMHPFYQY 408
DB 436 PNVCDMGKNGKAFIRFEADPMKNITVF 463
RESULT 36
Q9BPW0 HUMAN
ID Q9BPW0_HUMAN PRELIMINARY; PRT; 484 AA.
AC Q9BPW0;
DT 01-UTN-2001 (TrEMBLrel. 17, Created)
DT 01-UTN-2001 (TrEMBLrel. 17, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DB PPP5C protein (Fragment).
GN Name=PPP5C;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Muscle;
RX MEDLINE=22388237; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marinina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.R.,
RA Brownstein M.J., Ueda T.B., Tohliyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heiton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.C., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Maira M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Muscle;
RA Strauberg R.L.
RN [3] submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
QY 11 FLOKRYRHHQARERQRCNMQIFQNLLEYASEDDQALYKFENDLIKHPQAGRNQVQ 70
DB 93 YIKGYRRAASMAAGK-----FR-----ALRD-----YETVYKYPHDKAKMKYQ 135
QY 71 -----GSAHV--SVLDKDD---DIVSEFG--DIVNAKIELPFRKNHIDL 107
DB 136 ECKNLVKQKAFERAJAGBEHKRSYVDSDIDISMTEDBYSGPKLRDGVTTISFMKE----- 191
QY 108 LIDVRKKRGNRLHPFYVALILREAKSLKQLPNISPVSTANVQGVTVCGDLHGKLDLL 167
DB 192 LMQWTKDQK--NLHKRYAVQYILQYKLVSKLSTVETTLKETEKITVCGDHGQFYDIL 249
QY 168 VYLHKNGLPSSNPVYFENDFDPRGKGLFVLLLLSYLIAPNNAVFLNRGNHDSVNNAR 227
DB 250 NIFELNGLPSEBNPVLFNDFDPRGSPSLFVLTITFGRKLYPDHFLHNRGNHETDNMQ 309
QY 228 RYGFIREVSKYPRNHRKILAFIDEVYRMLPLGSVLSRVLIYVHG--SDSTSLDLK 285
DB 310 IYGFGEKVAKYT--AAMYELFSEVFEMPLPLAQCIKNGKVLIMHGLFSBQVITLDDIRK 366
QY 286 IDRGVYVILRPPLTDEBPLDKTEWQQLFDIWMSPQATMGCVPTLRAGVFGPDVTD 345
DB 367 IERNK-----QPP--DSQPM-----CDLWSPQONG--RSISKGVSCQFPDVTYK 410
QY 346 NQLQHRRLSYVLRSHCKRNGHEFMHDKITITFSASNYAIGSNKAYIRLN--NQLMHP 404
DB 411 AFLERNNDIYIIRSHSVDEGVIEIRHDKLITVFAPNVCDMGKNGKAYIRHIGSDILAPQ 470

QY 405 FVOY 408
Db 471 FHOF 474

RESULT 37
PPPS_HUMAN STANDARD; PRT; 499 AA.
ID PP5_HUMAN
AC P53041; Q16722;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Serine/threonine protein phosphatase 5 (EC 3.1.3.16) (PP5) (Protein
phosphatase T) (PP-T) (PPT).
GN Name=PP5C; Synonym=PPPS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Barchonoglires; Primates; Catarrhini; Homidae;
OC Homo.
NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE OF 7-499.
RX MEDLINE=95009929; PubMed=7925273;
RA Chen M.X., McPartlin A.E., Brown L., Chen Y.H., Barker H.M.,
Cohen P.T.W.;
RT "A novel human protein serine/threonine phosphatase, which possesses
four tetratricopeptide repeat motifs and localizes to the nucleus.";
RL EMBO J. 13:4278-4290(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 9-499.
RX MEDLINE=96115607; PubMed=8666404;
RA Yong W.H., Ueki K., Chou D., Reeves S.A., von Deimling A.,
Gusella J.F., Mohrenweiser H.W., Buckler A.J., Louis D.N.;
RT "Cloning of a highly conserved human protein serine-threonine
phosphatase gene from the glioma candidate region on chromosome
19q13.3.";
RL Genomics 29:533-536(1995).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15057824; DOI=10.1038/nature02399;
RA Grimwood J., Gordon L.A., Olsen A.S., Terry A., Schmutz J.,
Lamerdin J.E., Hellsten U., Goodstein D., Conrume O., Tran-Gyamfi M.,
Aerts A., Altherr M., Ashworth L., Bajorek E., Black S., Branscomb E.,
Caenepeel S., Carrano A.V., Cawley C., Chan Y.M., Christensen M.,
Cleland C.A., Copeland A., Dalin B., Dehal P., Denye M., Dettler J.C.,
Escobar J., Flowers D., Fotopoulos D., Garcia C., Georgescu A.M.,
Glavina T., Gomez M., Gonzales B., Groza M., Hammon N., Hawkins T.,
Hayda L., Ho I., Huang W., Israni S., Jett J., Kadner K., Kimball H.,
Kolayashli A., Laitonov V., Leem S.-H., Lopez F., Lou Y., Lowry S.,
Malfatti S., Martinez D., McCreedy P.M., Medina C., Morgan J.,
Nelson K., Nolan M., Ovcharenko I., Piltuck S., Pollard M.,
Popkie A.P., Predki P., Quan G., Ramirez L., Raab S., Retterer J.,
Rodriguez A., Rogers S., Salamov A., Salazar A., She X., Smith D.,
Slezak T., Solovyev V., Thayer N., Tice H., Tsai M., Ustaszewska A.,
Vo N., Wagner M., Wheeler J., Wu K., Xie G., Xie J., Yang J., Dubchak I.,
Parey T.S., DeLong P., Dickson M., Gordon D., Richter E.B.,
Pennacchio L.A., Richardson P., Stubbs L., Rokhsar D.S., Myers R.M.,
Rubin E.M., Lucas S.M.,
RT "The DNA sequence and biology of human chromosome 19.";
RL Nature 428:529-535(2004).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RX MEDLINE=22288257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
Klauser R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Datchenko L., Marusik K., Farmer A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Cassavani T.L., Scheetz T.E.,
Brownstein M.J., Uudin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahay J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalys D.B.,
RA Scherach A., Schein J.E., Jones S.J.M., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 1-37.
RC TISSUE=Fetal brain;
RX MEDLINE=96144708; PubMed=8561788; DOI=10.1006/bbrc.1996.0092;
RA Xu X., Lagercrantz J., Zickert P., Bajalica-Lagercrantz S.,
RA Zetterberg A.;
RT "Chromosomal localization and 5' sequence of the human protein
serine/threonine phosphatase 5' gene.";
RL Biochem. Biophys. Res. Commun. 218:514-517(1996).
RN [6]
RP INTERACTIONS WITH CDC16 AND CDC27.
RX MEDLINE=98070360; PubMed=9405394; DOI=10.1074/jbc.272.51.32011;
RA Ollendorff V., Donoghue D.J.;
RT "The serine/threonine phosphatase PP5 interacts with CDC16 and CDC27,
two tetratricopeptide repeat-containing subunits of the anaphase-
promoting complex.";
RL J. Biol. Chem. 272:32011-32018(1997).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.45 ANGSTROMS) OF 19-177.
RX MEDLINE=98151343; PubMed=9482716; DOI=10.1093/emboj/17.5.1192;
RA Das A.K., Cohen P.W., Barford D.;
RT "The structure of the tetratricopeptide repeats of protein phosphatase
5: implications for TPR-mediated protein-protein interactions.";
RL EMBO J. 17:1192-1199(1998).
CC -1- FUNCTION: May play a role in the regulation of RNA biogenesis
and/or mitosis. In vitro, dephosphorylates serine residues of
skeletal muscle phosphorylase and histone H1.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
phosphate.
CC -1- COFACTOR: Binds 1 iron ion per subunit (By similarity).
CC -1- COFACTOR: Binds 1 manganese ion per subunit (By similarity).
CC -1- SUBUNIT: Interacts with CDC16 and CDC27.
CC -1- SUBCELLULAR LOCATION: Nuclear; predominantly, but also present in
the cytoplasm.
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: Belongs to the PPP phosphatase family. PP-T subfamily.
CC -1- SIMILARITY: Contains 3 TPR repeats.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
DR EMBL; X89416; CAA61595.1; -; mRNA.
DR EMBL; U25174; AAB60384.1; -; mRNA.
DR EMBL; AC007193; AAD22669.1; -; Genomic DNA.
DR EMBL; BC001970; AA01970.1; -; mRNA.
DR EMBL; X92121; CAA63089.1; -; mRNA.
DR PIR; S52570; S52570.
DR PDB; 1AI7; X-ray; @=16-181.
DR PDB; 1S95; X-ray; A/B=169-499.
DR SMR; P53041; 23-499.
DR Ensembl; ENSG0000011485; Homo sapiens.
DR HGN; HGNC:9322; PP5C.
DR H-InvDB; HIX0015250; -.
DR MIM; 600658; -.
DR GO; GO:0005634; Cytoplasm; TAS.
DR GO; GO:004722; P:protein serine/threonine phosphatase activity; TAS.

```

DR GO; GO:0004871; F:signal transducer activity; IMP.
DR GO; GO:0007067; P:mitosis; TAS.
DR GO; GO:0004323; P:protein regulation of I-kappaB kinase/NF-kappaB; IMP.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR GO; GO:0006350; P:transcription; TAS.
DR InterPro; IPR004843; M-peptidase.
DR InterPro; IPR011236; P:peptase_5.
DR InterPro; IPR006186; T:phatase_apah.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR011990; TPR-like_helical.
DR Pfam; PF00149; Metallophos; 1.
DR Pfam; PF00515; TPR_1; 3.
DR PIRSF; PIRSF03096; P:peptase_5; 1.
DR PRINTS; PR00114; STPHPTAS.
DR ProDom; PD000252; T.phatase_apah; 1.
DR SMART; SM00156; PP2Ac; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS00125; SER_THR_PHOSPHATASE; 1.
DR PROSITE; PS50005; TPR; 3.
DR PROSITE; PS50293; TPR_REGION; 1.
DR 3D-structure; Hydroxylase; Iron; Manganese; Metal-binding; Nuclear protein; Protein phosphatase; Repeat; TPR repeat.
KM REPEAT; 28; 61; TPR 1.
FT REPEAT; 62; 95; TPR 2.
FT REPEAT; 96; 129; TPR 3.
FT REGION; 184; 499; Catalytic.
FT ACT SITE; 304; 304; Proton donor (By similarity).
FT METAL; 242; 242; Iron (By similarity).
FT METAL; 244; 244; Iron (By similarity).
FT METAL; 271; 271; Iron and manganese (By similarity).
FT METAL; 303; 303; Manganese (By similarity).
FT METAL; 352; 352; Manganese (By similarity).
FT METAL; 427; 427; Manganese (By similarity).
FT METAL; 403; 403; S -> T (in Ref. 2).
FT CONFLICT; 403; 403;
FT HELIX; 22; 40;
FT TURN; 41; 42;
FT HELIX; 44; 57;
FT TURN; 59; 60;
FT HELIX; 62; 74;
FT TURN; 75; 76;
FT HELIX; 78; 91;
FT TURN; 93; 94;
FT HELIX; 96; 108;
FT TURN; 109; 110;
FT HELIX; 112; 125;
FT TURN; 127; 128;
FT HELIX; 130; 164;
FT TURN; 168; 169;
SQ SEQUENCE 499 AA; 56879 MW; DB3B2090D8658BB3 CRC64;

Query Match 16.6%; Score 570; DB 1; Length 499;
Best Local Similarity 33.7%; Pred. No. 2,9e-26;
Matches 143; Conservative 71; Mismatches 142; Indels 68; Gaps 16;

```

```

QY 266 IDRGKYSILRPPLDSDPDLKTEWQIFDIMSNDPQATMGCVPTLTAGAGVIRGPDVTD 345
DB 369 IERNK-----QPP--DSGPM-----CDLIMSDPQPNQ--RSISKGVSCQFepdvtk 412
QY 346 NTLQRRRLSYVIRSHCKPENGHGFPMNDKILITIPASNYAIGSKNGAYIRL--NNQLMPH 404
DB 413 AFLERNNDYILIRSHVYAEIGYVAHGRCVTFSPAPNYCDQMGKASTYIHLQSGDLRPQ 472
QY 405 FVOY 408
DB 473 FHOQ 476

RESULT 38
Q53FR0_HUMAN
ID Q53FR0_HUMAN PRELIMINARY; PRT; 499 AA.
AC Q53FR0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Protein phosphatase 5, catalytic subunit variant (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC Homo.
NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=stomach mucosa;
RA Maruyama K., Sugano S.;
RT "Oligo-capping: a simple method to replace the cap structure of eucaryotic mRNAs with oligoribonucleotides.";
RL Gene 138:171-174(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=stomach mucosa;
RA Suzuki Y., Yoshitomo K., Maruyama K., Sugama A., Sugano S.;
RT "Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library.";
RL Gene 200:149-156(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=stomach mucosa;
RA Suzuki Y., Sugano S., Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.;
RT Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein + phosphate.
CC -!- SIMILARITY: Belongs to the PPP phosphatase family.
DR EMBL; AK232222; BAD96942.1; -; mRNA.
DR SMR; Q53FR0; 23-499.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA.
DR GO; GO:0006470; F:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR004843; M-peptidase.
DR InterPro; IPR006186; T.phatase_apah.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00149; Metallophos; 1.
DR Pfam; PF00515; TPR 1; 3.
DR PRINTS; PR00114; STPHPTAS.
DR ProDom; PD000252; T.phatase_apah; 1.
DR SMART; SM00156; PP2Ac; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS00125; SER_THR_PHOSPHATASE; UNKNOWN_1.
DR PROSITE; PS50005; TPR; 3.
DR PROSITE; PS50293; TPR_REGION; 1.
DR 3D-structure; Iron; TPR repeat.
FT NON TER 1
SQ SEQUENCE 499 AA; 56907 MW; DB5B20D85465C3F CRC64;

```

```

Query Match      16.6%; Score 570; DB 2; Length 499;
Match Local Similarity 33.7%; Pred.No.2.9e-26;
Matches 143; Conservative 71; PidentMatches 142; Indels 68; Gaps 16

QY 11 FLOKYYRHHQARERQRRCNMQIFONLEJASBQDAELIYKPFNDILIKHMPQAAGKRNQY 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 95 YIKGYRRRAASMAALGK-----FR-----AALRD-----YETVYKYPHDKAKMKYQ 137
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 71 -----GSAHV-SVLDDKD---DLVEFG--DIVNAKIELPIRKHIDL 107
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 138 ECKNIKOKAFERAIAGDEHKRSVDSIDIBSMTIEDFSGPKLBDKVTISFMKG----- 193
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 108 LIDVRKKRGRGNRLHFKRYALLIREAKSKLQKPNISSPVSTASQOYVTCGDLHGKLDLL 167
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 194 LMQWKDK--KLHKKCAVQILIVQKEVLSKSTLVEYTLTKETKXITVCGDTHGFPYDL 251
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 168 VLIHNGKPPSSSNPVYFNGDFVDYGRKGLEVLLLLSLYLAPFNVAVPLNRGNHDSVNA 227
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 252 NIFELNGLPSEINPFIIFNDQFVDRKSFVSVEVLLTFGRFLYLPDFHILNRHGRIDNMQ 311
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 228 RYGFIREVSKYPRNHKKILAFIDEVYKMLPGSVLNSRVLIIVHGCF--SDSTSLDIKS 285
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 312 IYGFEGEYKAKYTL--AQWELPSEVFEWLPLAQCIINGKVLIMHGGLPSEDDVTLDIRK 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 286 IDRGRYVSLAPRLTDBERPLDKTERMQQLFDIMKSPQATMGCVPTLTLAGACWFGPDVYT 345
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 369 IERN-----QPP--DSGPM-----CDLMSPPQPPONG--RSISKGVSCQFQPDVTK 412
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 346 NPLQHRHSYVYRSHCEKPNGHGFHMDKIIITIFASPNVYAIAGSKGAYIRL--NQQLMPH 404
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 413 AFLERNMDYIIRSHVKABGVETVAHGRCVTFSAFNYCDMGKASVYIHQSSDLRPQ 472
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 405 FVQY 408
   : :
DB 473 FHQF 476

RESULT 39
Q53XV2 HUMAN PRLIMINARY; PRT; 499 AA.
ID AC Q53XV2
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DB Protein phosphatase 5, catalytic subunit.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Butheria; Euarchontoglires; Primates; Catarrhini; Homindaes;
OC Homo.
OC NCBI_TaxID=9606;
OC (1)
RN NUCLEOTIDE SEQUENCE.
RA Kainine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length cDNAs in BD Creator (TM) System Donor
RT vector."
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2O) = a protein +
CC phosphate.
CC -1- SIMILARITY: Belongs to the PPP phosphatase family.
CC EMBL; BT007275; AAP5939.1; -; mRNA.
DR SRM; Q53XV2; 23-499.
DR Ensembl; ENSG00000011485; Homo sapiens.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR004843; M-peptidase.
DR InterPro; IPR006186; T:phosphatase.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR011990; TPR-like_helical.

```

| Query Match | 16.6% | Score 570; | DB 2; | Length 499; |
|--|------------------|--------------------|------------|-------------|
| Beet Local Similarity | 33.7% | Pred. No. 2.9e-26; | | |
| Matches 143; | Conservative 71; | Mismatches 142; | Indels 68; | Gaps 16; |
| DR Pfam: PF00149; Metallophos; 1. | | | | |
| DR Pfam: PF00515; TPR_1; 3 | | | | |
| DR PRINTS; PR00114; STPHPTASE. | | | | |
| DR Prodom; PD000252; T_phcse_apah; 1. | | | | |
| DR SMART; SMO0156; PP2AC; 1. | | | | |
| DR SMART; SMO0028; TPR; 3. | | | | |
| DR PROSITE; PS00125; SER_THR_PHOSPHATASE; UNKNOWN_1. | | | | |
| DR PROSITE; PS50005; TPR; 3. | | | | |
| DR PROSITE; PS50293; TPR_REGION; 1. | | | | |
| KW Hydrolase; Iron; Repeat; TPR repeat. | | | | |
| SEQUENCE 499 AA; 56879 MW; DB3B2090D8658B83 CRC64; | | | | |
| Query Match | 16.6% | Score 570; | DB 2; | Length 499; |
| Beet Local Similarity | 33.7% | Pred. No. 2.9e-26; | | |
| Matches 143; | Conservative 71; | Mismatches 142; | Indels 68; | Gaps 16; |
| DR Pfam: PF00149; Metallophos; 1. | | | | |
| DR Pfam: PF00515; TPR_1; 3 | | | | |
| DR PRINTS; PR00114; STPHPTASE. | | | | |
| DR Prodom; PD000252; T_phcse_apah; 1. | | | | |
| DR SMART; SMO0156; PP2AC; 1. | | | | |
| DR SMART; SMO0028; TPR; 3. | | | | |
| DR PROSITE; PS00125; SER_THR_PHOSPHATASE; UNKNOWN_1. | | | | |
| DR PROSITE; PS50005; TPR; 3. | | | | |
| DR PROSITE; PS50293; TPR_REGION; 1. | | | | |
| KW Hydrolase; Iron; Repeat; TPR repeat. | | | | |
| SEQUENCE 499 AA; 56879 MW; DB3B2090D8658B83 CRC64; | | | | |

